SEQUENCE LISTING

<110>  E. I. duPont de Nemours and Company, Inc.
       Cheng, Qiong
       Suh, Wonchul

<120>  Mutations Affecting Plasmid Copy Number

<130>  CL2029 US NA

<150>  US 60/434973
<151>  2002-12-20

<160>  28

<170>  PatentIn version 3.2

<210>  1
<211>  912
<212>  DNA
<213>  Pantoea stewartii


<220>
<221>  misc_feature
<222>  (1)..(3)
<223>  Alternative start code TTG instead of ATG used.

<400>  1
ttgacggtct gcgcaaaaaa acacgttcac cttactggca tttcggctga gcagttgctg      60

gctgatatcg atagccgcct tgatcagtta ctgccggttc agggtgagcg ggattgtgtg     120

ggtgccgcga tgcgtgaagg cacgctggca ccgggcaaac gtattcgtcc gatgctgctg     180

ttattaacag cgcgcgatct tggctgtgcg atcagtcacg ggggattact ggatttagcc     240

tgcgcggttg aaatggtgca tgctgcctcg ctgattctgg atgatatgcc ctgcatggac     300

gatgcgcaga tgcgtcgggg gcgtcccacc attcacacgc agtacggtga acatgtggcg     360

attctggcgg cggtcgcttt actcagcaaa gcgtttgggg tgattgccga ggctgaaggt     420

ctgacgccga tagccaaaac tcgcgcggtg tcggagctgt ccactgcgat tggcatgcag     480

ggtctggttc agggccagtt taaggacctc tcggaaggcg ataaaccccg cagcgccgat     540

gccatactgc taaccaatca gtttaaaacc agcacgctgt tttgcgcgtc aacgcaaatg     600

gcgtccattg cggccaacgc gtcctgcgaa gcgcgtgaga acctgcatcg tttctcgctc     660

gatctcggcc aggcctttca gttgcttgac gatcttaccg atggcatgac cgataccggc     720

aaagacatca atcaggatgc aggtaaatca acgctggtca atttattagg ctcaggcgcg     780

gtcgaagaac gcctgcgaca gcatttgcgc ctggccagtg aacacctttc gcgggcatgc     840

caaaacggcc attccaccac ccaacttttt attcaggcct ggtttgacaa aaaactcgct     900

gccgtcagtt aa                                                          912


<210>  2
<211>  303
<212>  PRT
<213>  Pantoea stewartii

<400>  2

```
Met Thr Val Cys Ala Lys Lys His Val His Leu Thr Gly Ile Ser Ala
1               5               10              15

Glu Gln Leu Leu Ala Asp Ile Asp Ser Arg Leu Asp Gln Leu Leu Pro
            20              25              30

Val Gln Gly Glu Arg Asp Cys Val Gly Ala Ala Met Arg Glu Gly Thr
        35              40              45

Leu Ala Pro Gly Lys Arg Ile Arg Pro Met Leu Leu Leu Leu Thr Ala
    50              55              60

Arg Asp Leu Gly Cys Ala Ile Ser His Gly Gly Leu Leu Asp Leu Ala
65              70              75              80

Cys Ala Val Glu Met Val His Ala Ala Ser Leu Ile Leu Asp Asp Met
            85              90              95

Pro Cys Met Asp Asp Ala Gln Met Arg Arg Gly Arg Pro Thr Ile His
            100             105             110

Thr Gln Tyr Gly Glu His Val Ala Ile Leu Ala Ala Val Ala Leu Leu
        115             120             125

Ser Lys Ala Phe Gly Val Ile Ala Glu Ala Glu Gly Leu Thr Pro Ile
    130             135             140

Ala Lys Thr Arg Ala Val Ser Glu Leu Ser Thr Ala Ile Gly Met Gln
145             150             155             160

Gly Leu Val Gln Gly Gln Phe Lys Asp Leu Ser Glu Gly Asp Lys Pro
            165             170             175

Arg Ser Ala Asp Ala Ile Leu Leu Thr Asn Gln Phe Lys Thr Ser Thr
            180             185             190

Leu Phe Cys Ala Ser Thr Gln Met Ala Ser Ile Ala Ala Asn Ala Ser
        195             200             205

Cys Glu Ala Arg Glu Asn Leu His Arg Phe Ser Leu Asp Leu Gly Gln
    210             215             220

Ala Phe Gln Leu Leu Asp Asp Leu Thr Asp Gly Met Thr Asp Thr Gly
225             230             235             240

Lys Asp Ile Asn Gln Asp Ala Gly Lys Ser Thr Leu Val Asn Leu Leu
            245             250             255

Gly Ser Gly Ala Val Glu Glu Arg Leu Arg Gln His Leu Arg Leu Ala
            260             265             270
```

```
Ser Glu His Leu Ser Ala Ala Cys Gln Asn Gly His Ser Thr Thr Gln
    275             280             285

Leu Phe Ile Gln Ala Trp Phe Asp Lys Lys Leu Ala Ala Val Ser
    290             295             300


<210>   3
<211>   1296
<212>   DNA
<213>   Pantoea stewartii


<220>
<221>   CDS
<222>   (1)..(1296)

<400>   3
atg agc cat ttt gcg gtg atc gca ccg ccc ttt ttc agc cat gtt cgc      48
Met Ser His Phe Ala Val Ile Ala Pro Pro Phe Phe Ser His Val Arg
1                 5               10              15

gct ctg caa aac ctt gct cag gaa tta gtg gcc cgc ggt cat cgt gtt      96
Ala Leu Gln Asn Leu Ala Gln Glu Leu Val Ala Arg Gly His Arg Val
            20              25              30

acg ttt ttt cag caa cat gac tgc aaa gcg ctg gta acg ggc agc gat     144
Thr Phe Phe Gln Gln His Asp Cys Lys Ala Leu Val Thr Gly Ser Asp
        35              40              45

atc gga ttc cag acc gtc gga ctg caa acg cat cct ccc ggt tcc tta     192
Ile Gly Phe Gln Thr Val Gly Leu Gln Thr His Pro Pro Gly Ser Leu
    50              55              60

tcg cac ctg ctg cac ctg gcc gcg cac cca ctc gga ccc tcg atg tta     240
Ser His Leu Leu His Leu Ala Ala His Pro Leu Gly Pro Ser Met Leu
65              70              75              80

cga ctg atc aat gaa atg gca cgt acc agc gat atg ctt tgc cgg gaa     288
Arg Leu Ile Asn Glu Met Ala Arg Thr Ser Asp Met Leu Cys Arg Glu
                85              90              95

ctg ccc gcc gct ttt cat gcg ttg cag ata gag ggc gtg atc gtt gat     336
Leu Pro Ala Ala Phe His Ala Leu Gln Ile Glu Gly Val Ile Val Asp
            100             105             110

caa atg gag ccg gca ggt gca gta gtc gca gaa gcg tca ggt ctg ccg     384
Gln Met Glu Pro Ala Gly Ala Val Val Ala Glu Ala Ser Gly Leu Pro
            115             120             125

ttt gtt tcg gtg gcc tgc gcg ctg ccg ctc aac cgc gaa ccg ggt ttg     432
Phe Val Ser Val Ala Cys Ala Leu Pro Leu Asn Arg Glu Pro Gly Leu
    130             135             140

cct ctg gcg gtg atg cct ttc gag tac ggc acc agc gat gcg gct cgg     480
Pro Leu Ala Val Met Pro Phe Glu Tyr Gly Thr Ser Asp Ala Ala Arg
145             150             155             160

gaa cgc tat acc acc agc gaa aaa att tat gac tgg ctg atg cga cgt     528
Glu Arg Tyr Thr Thr Ser Glu Lys Ile Tyr Asp Trp Leu Met Arg Arg
                165             170             175

cac gat cgt gtg atc gcg cat cat gca tgc aga atg ggt tta gcc ccg     576
His Asp Arg Val Ile Ala His His Ala Cys Arg Met Gly Leu Ala Pro
            180             185             190
```

```
cgt gaa aaa ctg cat cat tgt ttt tct cca ctg gca caa atc agc cag      624
Arg Glu Lys Leu His His Cys Phe Ser Pro Leu Ala Gln Ile Ser Gln
        195                 200                 205

ttg atc ccc gaa ctg gat ttt ccc cgc aaa gcg ctg cca gac tgc ttt      672
Leu Ile Pro Glu Leu Asp Phe Pro Arg Lys Ala Leu Pro Asp Cys Phe
        210                 215                 220

cat gcg gtt gga ccg tta cgg caa ccc cag ggg acg ccg ggg tca tca      720
His Ala Val Gly Pro Leu Arg Gln Pro Gln Gly Thr Pro Gly Ser Ser
225                 230                 235                 240

act tct tat ttt ccg tcc ccg gac aaa ccc cgt att ttt gcc tcg ctg      768
Thr Ser Tyr Phe Pro Ser Pro Asp Lys Pro Arg Ile Phe Ala Ser Leu
                245                 250                 255

ggc acc ctg cag gga cat cgt tat ggc ctg ttc agg acc atc gcc aaa      816
Gly Thr Leu Gln Gly His Arg Tyr Gly Leu Phe Arg Thr Ile Ala Lys
                260                 265                 270

gcc tgc gaa gag gtg gat gcg cag tta ctg ttg gca cac tgt ggc ggc      864
Ala Cys Glu Glu Val Asp Ala Gln Leu Leu Leu Ala His Cys Gly Gly
            275                 280                 285

ctc tca gcc acg cag gca ggt gaa ctg gcc cgg ggc ggg gac att cag      912
Leu Ser Ala Thr Gln Ala Gly Glu Leu Ala Arg Gly Gly Asp Ile Gln
        290                 295                 300

gtt gtg gat ttt gcc gat caa tcc gca gca ctt tca cag gca cag ttg      960
Val Val Asp Phe Ala Asp Gln Ser Ala Ala Leu Ser Gln Ala Gln Leu
305                 310                 315                 320

aca atc aca cat ggt ggg atg aat acg gta ctg gac gct att gct tcc     1008
Thr Ile Thr His Gly Gly Met Asn Thr Val Leu Asp Ala Ile Ala Ser
                325                 330                 335

cgc aca ccg cta ctg gcg ctg ccg ctg gca ttt gat caa cct ggc gtg     1056
Arg Thr Pro Leu Leu Ala Leu Pro Leu Ala Phe Asp Gln Pro Gly Val
                340                 345                 350

gca tca cga att gtt tat cat ggc atc ggc aag cgt gcg tct cgg ttt     1104
Ala Ser Arg Ile Val Tyr His Gly Ile Gly Lys Arg Ala Ser Arg Phe
        355                 360                 365

act acc agc cat gcg ctg gcg cgg cag att cga tcg ctg ctg act aac     1152
Thr Thr Ser His Ala Leu Ala Arg Gln Ile Arg Ser Leu Leu Thr Asn
        370                 375                 380

acc gat tac ccg cag cgt atg aca aaa att cag gcc gca ttg cgt ctg     1200
Thr Asp Tyr Pro Gln Arg Met Thr Lys Ile Gln Ala Ala Leu Arg Leu
385                 390                 395                 400

gca ggc ggc aca cca gcc gcc gcc gat att gtt gaa cag gcg atg cgg     1248
Ala Gly Gly Thr Pro Ala Ala Ala Asp Ile Val Glu Gln Ala Met Arg
                405                 410                 415

acc tgt cag cca gta ctc agt ggg cag gat tat gca acc gca cta tga     1296
Thr Cys Gln Pro Val Leu Ser Gly Gln Asp Tyr Ala Thr Ala Leu
                420                 425                 430
```

<210>   4
<211>   431
<212>   PRT
<213>   Pantoea stewartii

<400>   4

```
Met Ser His Phe Ala Val Ile Ala Pro Pro Phe Phe Ser His Val Arg
1               5               10              15

Ala Leu Gln Asn Leu Ala Gln Glu Leu Val Ala Arg Gly His Arg Val
            20              25              30

Thr Phe Phe Gln Gln His Asp Cys Lys Ala Leu Val Thr Gly Ser Asp
        35              40              45

Ile Gly Phe Gln Thr Val Gly Leu Gln Thr His Pro Pro Gly Ser Leu
    50              55              60

Ser His Leu Leu His Leu Ala Ala His Pro Leu Gly Pro Ser Met Leu
65              70              75              80

Arg Leu Ile Asn Glu Met Ala Arg Thr Ser Asp Met Leu Cys Arg Glu
            85              90              95

Leu Pro Ala Ala Phe His Ala Leu Gln Ile Glu Gly Val Ile Val Asp
            100             105             110

Gln Met Glu Pro Ala Gly Ala Val Val Ala Glu Ala Ser Gly Leu Pro
        115             120             125

Phe Val Ser Val Ala Cys Ala Leu Pro Leu Asn Arg Glu Pro Gly Leu
    130             135             140

Pro Leu Ala Val Met Pro Phe Glu Tyr Gly Thr Ser Asp Ala Ala Arg
145             150             155             160

Glu Arg Tyr Thr Thr Ser Glu Lys Ile Tyr Asp Trp Leu Met Arg Arg
            165             170             175

His Asp Arg Val Ile Ala His His Ala Cys Arg Met Gly Leu Ala Pro
            180             185             190

Arg Glu Lys Leu His His Cys Phe Ser Pro Leu Ala Gln Ile Ser Gln
        195             200             205

Leu Ile Pro Glu Leu Asp Phe Pro Arg Lys Ala Leu Pro Asp Cys Phe
    210             215             220

His Ala Val Gly Pro Leu Arg Gln Pro Gln Gly Thr Pro Gly Ser Ser
225             230             235             240

Thr Ser Tyr Phe Pro Ser Pro Asp Lys Pro Arg Ile Phe Ala Ser Leu
            245             250             255

Gly Thr Leu Gln Gly His Arg Tyr Gly Leu Phe Arg Thr Ile Ala Lys
            260             265             270
```

```
Ala Cys Glu Glu Val Asp Ala Gln Leu Leu Leu Ala His Cys Gly Gly
        275             280                 285


Leu Ser Ala Thr Gln Ala Gly Glu Leu Ala Arg Gly Gly Asp Ile Gln
    290             295                 300


Val Val Asp Phe Ala Asp Gln Ser Ala Ala Leu Ser Gln Ala Gln Leu
305             310                 315                 320


Thr Ile Thr His Gly Gly Met Asn Thr Val Leu Asp Ala Ile Ala Ser
            325                 330                 335


Arg Thr Pro Leu Leu Ala Leu Pro Leu Ala Phe Asp Gln Pro Gly Val
            340                 345                 350


Ala Ser Arg Ile Val Tyr His Gly Ile Gly Lys Arg Ala Ser Arg Phe
            355                 360                 365


Thr Thr Ser His Ala Leu Ala Arg Gln Ile Arg Ser Leu Leu Thr Asn
    370                 375                 380


Thr Asp Tyr Pro Gln Arg Met Thr Lys Ile Gln Ala Ala Leu Arg Leu
385             390                 395                 400


Ala Gly Gly Thr Pro Ala Ala Ala Asp Ile Val Glu Gln Ala Met Arg
            405                 410                 415


Thr Cys Gln Pro Val Leu Ser Gly Gln Asp Tyr Ala Thr Ala Leu
            420                 425                 430
```

<210>   5
<211>   1149
<212>   DNA
<213>   Pantoea stewartii


<220>
<221>   CDS
<222>   (1)..(1149)

<400>   5
```
atg caa ccg cac tat gat ctc att ctg gtc ggt gcc ggt ctg gct aat     48
Met Gln Pro His Tyr Asp Leu Ile Leu Val Gly Ala Gly Leu Ala Asn
1               5                   10                  15

ggc ctt atc gcg ctc cgg ctt cag caa cag cat ccg gat atg cgg atc     96
Gly Leu Ile Ala Leu Arg Leu Gln Gln Gln His Pro Asp Met Arg Ile
            20                  25                  30

ttg ctt att gag gcg ggt cct gag gcg gga ggg aac cat acc tgg tcc    144
Leu Leu Ile Glu Ala Gly Pro Glu Ala Gly Gly Asn His Thr Trp Ser
            35                  40                  45

ttt cac gaa gag gat tta acg ctg aat cag cat cgc tgg ata gcg ccg    192
Phe His Glu Glu Asp Leu Thr Leu Asn Gln His Arg Trp Ile Ala Pro
    50                  55                  60

ctt gtg gtc cat cac tgg ccc gac tac cag gtt cgt ttc ccc caa cgc    240
```

```
Leu Val Val His His Trp Pro Asp Tyr Gln Val Arg Phe Pro Gln Arg
65              70              75              80

cgt cgc cat gtg aac agt ggc tac tac tgc gtg acc tcc cgg cat ttc      288
Arg Arg His Val Asn Ser Gly Tyr Tyr Cys Val Thr Ser Arg His Phe
                85              90              95

gcc ggg ata ctc cgg caa cag ttt gga caa cat tta tgg ctg cat acc      336
Ala Gly Ile Leu Arg Gln Gln Phe Gly Gln His Leu Trp Leu His Thr
            100             105             110

gcg gtt tca gcc gtt cat gct gaa tcg gtc cag tta gcg gat ggc cgg      384
Ala Val Ser Ala Val His Ala Glu Ser Val Gln Leu Ala Asp Gly Arg
            115             120             125

att att cat gcc agt aca gtg atc gac gga cgg ggt tac acg cct gat      432
Ile Ile His Ala Ser Thr Val Ile Asp Gly Arg Gly Tyr Thr Pro Asp
            130             135             140

tct gca cta cgc gta gga ttc cag gca ttt atc ggt cag gag tgg caa      480
Ser Ala Leu Arg Val Gly Phe Gln Ala Phe Ile Gly Gln Glu Trp Gln
145             150             155             160

ctg agc gcg ccg cat ggt tta tcg tca ccg att atc atg gat gcg acg      528
Leu Ser Ala Pro His Gly Leu Ser Ser Pro Ile Ile Met Asp Ala Thr
                165             170             175

gtc gat cag caa aat ggc tac cgc ttt gtt tat acc ctg ccg ctt tcc      576
Val Asp Gln Gln Asn Gly Tyr Arg Phe Val Tyr Thr Leu Pro Leu Ser
            180             185             190

gca acc gca ctg ctg atc gaa gac aca cac tac att gac aag gct aat      624
Ala Thr Ala Leu Leu Ile Glu Asp Thr His Tyr Ile Asp Lys Ala Asn
            195             200             205

ctt cag gcc gaa cgg gcg cgt cag aac att cgc gat tat gct gcg cga      672
Leu Gln Ala Glu Arg Ala Arg Gln Asn Ile Arg Asp Tyr Ala Ala Arg
            210             215             220

cag ggt tgg ccg tta cag acg ttg ctg cgg gaa gaa cag ggt gca ttg      720
Gln Gly Trp Pro Leu Gln Thr Leu Leu Arg Glu Glu Gln Gly Ala Leu
225             230             235             240

ccc att acg tta acg ggc gat aat cgt cag ttt tgg caa cag caa ccg      768
Pro Ile Thr Leu Thr Gly Asp Asn Arg Gln Phe Trp Gln Gln Gln Pro
                245             250             255

caa gcc tgt agc gga tta cgc gcc ggg ctg ttt cat ccg aca acc ggc      816
Gln Ala Cys Ser Gly Leu Arg Ala Gly Leu Phe His Pro Thr Thr Gly
            260             265             270

tac tcc cta ccg ctc gcg gtg gcg ctg gcc gat cgt ctc agc gcg ctg      864
Tyr Ser Leu Pro Leu Ala Val Ala Leu Ala Asp Arg Leu Ser Ala Leu
            275             280             285

gat gtg ttt acc tct tcc tct gtt cac cag acg att gct cac ttt gcc      912
Asp Val Phe Thr Ser Ser Ser Val His Gln Thr Ile Ala His Phe Ala
            290             295             300

cag caa cgt tgg cag caa cag ggg ttt ttc cgc atg ctg aat cgc atg      960
Gln Gln Arg Trp Gln Gln Gln Gly Phe Phe Arg Met Leu Asn Arg Met
305             310             315             320

ttg ttt tta gcc gga ccg gcc gag tca cgc tgg cgt gtg atg cag cgt      1008
Leu Phe Leu Ala Gly Pro Ala Glu Ser Arg Trp Arg Val Met Gln Arg
            325             330             335

ttc tat ggc tta ccc gag gat ttg att gcc cgc ttt tat gcg gga aaa      1056
Phe Tyr Gly Leu Pro Glu Asp Leu Ile Ala Arg Phe Tyr Ala Gly Lys
```

```
Phe Tyr Gly Leu Pro Glu Asp Leu Ile Ala Arg Phe Tyr Ala Gly Lys
        340                 345                 350

ctc acc gtg acc gat cgg cta cgc att ctg agc ggc aag ccg ccc gtt      1104
Leu Thr Val Thr Asp Arg Leu Arg Ile Leu Ser Gly Lys Pro Pro Val
        355                 360                 365

ccc gtt ttc gcg gca ttg cag gca att atg acg act cat cgt tga          1149
Pro Val Phe Ala Ala Leu Gln Ala Ile Met Thr Thr His Arg
        370                 375                 380
```

<210>  6
<211>  382
<212>  PRT
<213>  Pantoea stewartii

<400>  6

```
Met Gln Pro His Tyr Asp Leu Ile Leu Val Gly Ala Gly Leu Ala Asn
1                   5                  10                  15

Gly Leu Ile Ala Leu Arg Leu Gln Gln Gln His Pro Asp Met Arg Ile
            20                  25                  30

Leu Leu Ile Glu Ala Gly Pro Glu Ala Gly Gly Asn His Thr Trp Ser
        35                  40                  45

Phe His Glu Glu Asp Leu Thr Leu Asn Gln His Arg Trp Ile Ala Pro
    50                  55                  60

Leu Val Val His His Trp Pro Asp Tyr Gln Val Arg Phe Pro Gln Arg
65                  70                  75                  80

Arg Arg His Val Asn Ser Gly Tyr Tyr Cys Val Thr Ser Arg His Phe
                85                  90                  95

Ala Gly Ile Leu Arg Gln Gln Phe Gly Gln His Leu Trp Leu His Thr
            100                 105                 110

Ala Val Ser Ala Val His Ala Glu Ser Val Gln Leu Ala Asp Gly Arg
        115                 120                 125

Ile Ile His Ala Ser Thr Val Ile Asp Gly Arg Gly Tyr Thr Pro Asp
    130                 135                 140

Ser Ala Leu Arg Val Gly Phe Gln Ala Phe Ile Gly Gln Glu Trp Gln
145                 150                 155                 160

Leu Ser Ala Pro His Gly Leu Ser Ser Pro Ile Ile Met Asp Ala Thr
                165                 170                 175

Val Asp Gln Gln Asn Gly Tyr Arg Phe Val Tyr Thr Leu Pro Leu Ser
            180                 185                 190

Ala Thr Ala Leu Leu Ile Glu Asp Thr His Tyr Ile Asp Lys Ala Asn
```
Page 8

```
            195                    200                    205
```

```
Leu Gln Ala Glu Arg Ala Arg Gln Asn Ile Arg Asp Tyr Ala Ala Arg
    210             215                 220
```

```
Gln Gly Trp Pro Leu Gln Thr Leu Leu Arg Glu Glu Gln Gly Ala Leu
225             230                 235                 240
```

```
Pro Ile Thr Leu Thr Gly Asp Asn Arg Gln Phe Trp Gln Gln Gln Pro
            245                 250                 255
```

```
Gln Ala Cys Ser Gly Leu Arg Ala Gly Leu Phe His Pro Thr Thr Gly
            260                 265                 270
```

```
Tyr Ser Leu Pro Leu Ala Val Ala Leu Ala Asp Arg Leu Ser Ala Leu
        275                 280                 285
```

```
Asp Val Phe Thr Ser Ser Ser Val His Gln Thr Ile Ala His Phe Ala
    290                 295                 300
```

```
Gln Gln Arg Trp Gln Gln Gln Gly Phe Phe Arg Met Leu Asn Arg Met
305             310                 315                 320
```

```
Leu Phe Leu Ala Gly Pro Ala Glu Ser Arg Trp Arg Val Met Gln Arg
            325                 330                 335
```

```
Phe Tyr Gly Leu Pro Glu Asp Leu Ile Ala Arg Phe Tyr Ala Gly Lys
            340                 345                 350
```

```
Leu Thr Val Thr Asp Arg Leu Arg Ile Leu Ser Gly Lys Pro Pro Val
        355                 360                 365
```

```
Pro Val Phe Ala Ala Leu Gln Ala Ile Met Thr Thr His Arg
    370                 375                 380
```

<210> 7
<211> 1479
<212> DNA
<213> Pantoea stewartii

<220>
<221> CDS
<222> (1)..(1479)

<400> 7
```
atg aaa cca act acg gta att ggt gcg ggc ttt ggt ggc ctg gca ctg    48
Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu
1               5                   10                  15
```

```
gca att cgt tta cag gcc gca ggt att cct gtt ttg ctg ctt gag cag    96
Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val Leu Leu Leu Glu Gln
                20                  25                  30
```

```
cgc gac aag ccg ggt ggc cgg gct tat gtt tat cag gag cag ggc ttt   144
Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr Gln Glu Gln Gly Phe
```

```
act ttt gat gca ggc cct acc gtt atc acc gat ccc agc gcg att gaa        192
Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Ser Ala Ile Glu
    50              55                  60

gaa ctg ttt gct ctg gcc ggt aaa cag ctt aag gat tac gtc gag ctg        240
Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys Asp Tyr Val Glu Leu
65              70                  75                  80

ttg ccg gtc acg ccg ttt tat cgc ctg tgc tgg gag tcc ggc aag gtc        288
Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Val
                85                  90                  95

ttc aat tac gat aac gac cag gcc cag tta gaa gcg cag ata cag cag        336
Phe Asn Tyr Asp Asn Asp Gln Ala Gln Leu Glu Ala Gln Ile Gln Gln
                100                 105                 110

ttt aat ccg cgc gat gtt gcg ggt tat cga gcg ttc ctt gac tat tcg        384
Phe Asn Pro Arg Asp Val Ala Gly Tyr Arg Ala Phe Leu Asp Tyr Ser
            115                 120                 125

cgt gcc gta ttc aat gag ggc tat ctg aag ctc ggc act gtg cct ttt        432
Arg Ala Val Phe Asn Glu Gly Tyr Leu Lys Leu Gly Thr Val Pro Phe
    130                 135                 140

tta tcg ttc aaa gac atg ctt cgg gcc gcg ccc cag ttg gca aag ctg        480
Leu Ser Phe Lys Asp Met Leu Arg Ala Ala Pro Gln Leu Ala Lys Leu
145                 150                 155                 160

cag gca tgg cgc agc gtt tac agt aaa gtt gcc ggc tac att gag gat        528
Gln Ala Trp Arg Ser Val Tyr Ser Lys Val Ala Gly Tyr Ile Glu Asp
                165                 170                 175

gag cat ctt cgg cag gcg ttt tct ttt cac tcg ctc tta gtg ggg ggg        576
Glu His Leu Arg Gln Ala Phe Ser Phe His Ser Leu Leu Val Gly Gly
            180                 185                 190

aat ccg ttt gca acc tcg tcc att tat acg ctg att cac gcg tta gaa        624
Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu Ile His Ala Leu Glu
            195                 200                 205

cgg gaa tgg ggc gtc tgg ttt cca cgc ggt gga acc ggt gcg ctg gtc        672
Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly Thr Gly Ala Leu Val
    210                 215                 220

aat ggc atg atc aag ctg ttt cag gat ctg ggc ggc gaa gtc gtg ctt        720
Asn Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Gly Glu Val Val Leu
225                 230                 235                 240

aac gcc cgg gtc agt cat atg gaa acc gtt ggg gac aag att cag gcc        768
Asn Ala Arg Val Ser His Met Glu Thr Val Gly Asp Lys Ile Gln Ala
                245                 250                 255

gtg cag ttg gaa gac ggc aga cgg ttt gaa acc tgc gcg gtg gcg tcg        816
Val Gln Leu Glu Asp Gly Arg Arg Phe Glu Thr Cys Ala Val Ala Ser
            260                 265                 270

aac gct gat gtt gta cat acc tat cgc gat ctg ctg tct cag cat ccc        864
Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro
            275                 280                 285

gca gcc gct aag cag gcg aaa aaa ctg caa tcc aag cgt atg agt aac        912
Ala Ala Ala Lys Gln Ala Lys Lys Leu Gln Ser Lys Arg Met Ser Asn
    290                 295                 300

tca ctg ttt gta ctc tat ttt ggt ctc aac cat cat cac gat caa ctc        960
Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu
```

```
305                         310                         315                         320
gcc cat cat acc gtc tgt ttt ggg cca cgc tac cgt gaa ctg att cac        1008
Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile His
                325                         330                 335

gaa att ttt aac cat gat ggt ctg gct gag gat ttt tcg ctt tat tta        1056
Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu
                340                         345                 350

cac gca cct tgt gtc acg gat ccg tca ctg gca ccg gaa ggg tgc ggc        1104
His Ala Pro Cys Val Thr Asp Pro Ser Leu Ala Pro Glu Gly Cys Gly
            355                         360                 365

agc tat tat gtg ctg gcg cct gtt cca cac tta ggc acg gcg aac ctc        1152
Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu
        370                         375                 380

gac tgg gcg gta gaa gga ccc cga ctg cgc gat cgt att ttt gac tac        1200
Asp Trp Ala Val Glu Gly Pro Arg Leu Arg Asp Arg Ile Phe Asp Tyr
385                         390                 395                 400

ctt gag caa cat tac atg cct ggc ttg cga agc cag ttg gtg acg cac        1248
Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His
                405                         410                 415

cgt atg ttt acg ccg ttc gat ttc cgc gac gag ctc aat gcc tgg caa        1296
Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Glu Leu Asn Ala Trp Gln
                420                         425                 430

ggt tcg gcc ttc tcg gtt gaa cct att ctg acc cag agc gcc tgg ttc        1344
Gly Ser Ala Phe Ser Val Glu Pro Ile Leu Thr Gln Ser Ala Trp Phe
            435                         440                 445

cga cca cat aac cgc gat aag cac att gat aat ctt tat ctg gtt ggc        1392
Arg Pro His Asn Arg Asp Lys His Ile Asp Asn Leu Tyr Leu Val Gly
        450                         455                 460

gca ggc acc cat cct ggc gcg ggc att ccc ggc gta atc ggc tcg gcg        1440
Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala
465                         470                 475                 480

aag gcg acg gca ggc tta atg ctg gag gac ctg att tga                    1479
Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile
                485                         490
```

```
<210>   8
<211>   492
<212>   PRT
<213>   Pantoea stewartii

<400>   8

Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu
1                   5                   10                  15


Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val Leu Leu Leu Glu Gln
                20                  25                  30


Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr Gln Glu Gln Gly Phe
            35                  40                  45


Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Ser Ala Ile Glu
        50                  55                  60
```

Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys Asp Tyr Val Glu Leu
65              70              75                      80

Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Val
                85              90                      95

Phe Asn Tyr Asp Asn Asp Gln Ala Gln Leu Glu Ala Gln Ile Gln Gln
            100             105                     110

Phe Asn Pro Arg Asp Val Ala Gly Tyr Arg Ala Phe Leu Asp Tyr Ser
            115             120                     125

Arg Ala Val Phe Asn Glu Gly Tyr Leu Lys Leu Gly Thr Val Pro Phe
    130             135                 140

Leu Ser Phe Lys Asp Met Leu Arg Ala Ala Pro Gln Leu Ala Lys Leu
145             150                 155                 160

Gln Ala Trp Arg Ser Val Tyr Ser Lys Val Ala Gly Tyr Ile Glu Asp
            165             170                     175

Glu His Leu Arg Gln Ala Phe Ser Phe His Ser Leu Leu Val Gly Gly
            180             185                     190

Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu Ile His Ala Leu Glu
        195             200                 205

Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly Thr Gly Ala Leu Val
    210             215                 220

Asn Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Gly Glu Val Val Leu
225             230                 235                 240

Asn Ala Arg Val Ser His Met Glu Thr Val Gly Asp Lys Ile Gln Ala
            245             250                     255

Val Gln Leu Glu Asp Gly Arg Arg Phe Glu Thr Cys Ala Val Ala Ser
            260             265                     270

Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro
            275             280                     285

Ala Ala Ala Lys Gln Ala Lys Lys Leu Gln Ser Lys Arg Met Ser Asn
    290             295                 300

Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu
305             310                 315                 320

Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile His
            325             330                     335

Page 12

```
Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu
        340             345             350


His Ala Pro Cys Val Thr Asp Pro Ser Leu Ala Pro Glu Gly Cys Gly
        355             360             365


Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu
    370             375             380


Asp Trp Ala Val Glu Gly Pro Arg Leu Arg Asp Arg Ile Phe Asp Tyr
385             390             395             400


Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His
            405             410             415


Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Glu Leu Asn Ala Trp Gln
        420             425             430


Gly Ser Ala Phe Ser Val Glu Pro Ile Leu Thr Gln Ser Ala Trp Phe
        435             440             445


Arg Pro His Asn Arg Asp Lys His Ile Asp Asn Leu Tyr Leu Val Gly
    450             455             460


Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala
465             470             475             480


Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile
            485             490
```

<210> 9
<211> 891
<212> DNA
<213> Pantoea stewartii


<220>
<221> CDS
<222> (1)..(891)

<400> 9
```
atg gcg gtt ggc tcg aaa agc ttt gcg act gca tcg acg ctt ttc gac      48
Met Ala Val Gly Ser Lys Ser Phe Ala Thr Ala Ser Thr Leu Phe Asp
1               5               10              15

gcc aaa acc cgt cgc agc gtg ctg atg ctt tac gca tgg tgc cgc cac      96
Ala Lys Thr Arg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His
            20              25              30

tgc gac gac gtc att gac gat caa aca ctg ggc ttt cat gcc gac cag     144
Cys Asp Asp Val Ile Asp Asp Gln Thr Leu Gly Phe His Ala Asp Gln
        35              40              45

ccc tct tcg cag atg cct gag cag cgc ctg cag cag ctt gaa atg aaa     192
Pro Ser Ser Gln Met Pro Glu Gln Arg Leu Gln Gln Leu Glu Met Lys
    50              55              60
```

```
acg cgt cag gcc tac gcc ggt tcg caa atg cac gag ccc gct ttt gcc      240
Thr Arg Gln Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala
65              70                  75                  80

gcg ttt cag gag gtc gcg atg gcg cat gat atc gct ccc gcc tac gcg      288
Ala Phe Gln Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala
                85                  90                  95

ttc gac cat ctg gaa ggt ttt gcc atg gat gtg cgc gaa acg cgc tac      336
Phe Asp His Leu Glu Gly Phe Ala Met Asp Val Arg Glu Thr Arg Tyr
            100                 105                 110

ctg aca ctg gac gat acg ctg cgt tat tgc tat cac gtc gcc ggt gtt      384
Leu Thr Leu Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val
            115                 120                 125

gtg ggc ctg atg atg gcg caa att atg ggc gtt cgc gat aac gcc acg      432
Val Gly Leu Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr
        130                 135                 140

ctc gat cgc gcc tgc gat ctc ggg ctg gct ttc cag ttg acc aac att      480
Leu Asp Arg Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile
145             150                 155                 160

gcg cgt gat att gtc gac gat gct cag gtg ggc cgc tgt tat ctg cct      528
Ala Arg Asp Ile Val Asp Asp Ala Gln Val Gly Arg Cys Tyr Leu Pro
            165                 170                 175

gaa agc tgg ctg gaa gag gaa gga ctg acg aaa gcg aat tat gct gcg      576
Glu Ser Trp Leu Glu Glu Glu Gly Leu Thr Lys Ala Asn Tyr Ala Ala
            180                 185                 190

cca gaa aac cgg cag gcc tta agc cgt atc gcc ggg cga ctg gta cgg      624
Pro Glu Asn Arg Gln Ala Leu Ser Arg Ile Ala Gly Arg Leu Val Arg
            195                 200                 205

gaa gcg gaa ccc tat tac gta tca tca atg gcc ggt ctg gca caa tta      672
Glu Ala Glu Pro Tyr Tyr Val Ser Ser Met Ala Gly Leu Ala Gln Leu
        210                 215                 220

ccc tta cgc tcg gcc tgg gcc atc gcg aca gcg aag cag gtg tac cgt      720
Pro Leu Arg Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg
225             230                 235                 240

aaa att ggc gtg aaa gtt gaa cag gcc ggt aag cag gcc tgg gat cat      768
Lys Ile Gly Val Lys Val Glu Gln Ala Gly Lys Gln Ala Trp Asp His
            245                 250                 255

cgc cag tcc acg tcc acc gcc gaa aaa tta acg ctt ttg ctg acg gca      816
Arg Gln Ser Thr Ser Thr Ala Glu Lys Leu Thr Leu Leu Leu Thr Ala
            260                 265                 270

tcc ggt cag gca gtt act tcc cgg atg aag acg tat cca ccc cgt cct      864
Ser Gly Gln Ala Val Thr Ser Arg Met Lys Thr Tyr Pro Pro Arg Pro
            275                 280                 285

gct cat ctc tgg cag cgc ccg atc tag                                  891
Ala His Leu Trp Gln Arg Pro Ile
290                 295
```

<210> 10
<211> 296
<212> PRT
<213> Pantoea stewartii

<400> 10

```
Met Ala Val Gly Ser Lys Ser Phe Ala Thr Ala Ser Thr Leu Phe Asp
1               5               10              15

Ala Lys Thr Arg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His
            20              25              30

Cys Asp Asp Val Ile Asp Asp Gln Thr Leu Gly Phe His Ala Asp Gln
        35              40              45

Pro Ser Ser Gln Met Pro Glu Gln Arg Leu Gln Gln Leu Glu Met Lys
    50              55              60

Thr Arg Gln Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala
65              70              75              80

Ala Phe Gln Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala
            85              90              95

Phe Asp His Leu Glu Gly Phe Ala Met Asp Val Arg Glu Thr Arg Tyr
            100             105             110

Leu Thr Leu Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val
        115             120             125

Val Gly Leu Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr
    130             135             140

Leu Asp Arg Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile
145             150             155             160

Ala Arg Asp Ile Val Asp Asp Ala Gln Val Gly Arg Cys Tyr Leu Pro
            165             170             175

Glu Ser Trp Leu Glu Glu Glu Gly Leu Thr Lys Ala Asn Tyr Ala Ala
            180             185             190

Pro Glu Asn Arg Gln Ala Leu Ser Arg Ile Ala Gly Arg Leu Val Arg
        195             200             205

Glu Ala Glu Pro Tyr Tyr Val Ser Ser Met Ala Gly Leu Ala Gln Leu
    210             215             220

Pro Leu Arg Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg
225             230             235             240

Lys Ile Gly Val Lys Val Glu Gln Ala Gly Lys Gln Ala Trp Asp His
            245             250             255

Arg Gln Ser Thr Ser Thr Ala Glu Lys Leu Thr Leu Leu Leu Thr Ala
        260             265             270
```

Page 15

```
Ser Gly Gln Ala Val Thr Ser Arg Met Lys Thr Tyr Pro Pro Arg Pro
        275                 280             285


Ala His Leu Trp Gln Arg Pro Ile
    290                 295


<210>   11
<211>   528
<212>   DNA
<213>   Pantoea stewartii


<220>
<221>   CDS
<222>   (1)..(528)

<400>   11
atg ttg tgg att tgg aat gcc ctg atc gtg ttt gtc acc gtg gtc ggc      48
Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Val Gly
1               5                   10                  15

atg gaa gtg gtt gct gca ctg gca cat aaa tac atc atg cac ggc tgg      96
Met Glu Val Val Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
            20                  25                  30

ggt tgg ggc tgg cat ctt tca cat cat gaa ccg cgt aaa ggc gca ttt     144
Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
        35                  40                  45

gaa gtt aac gat ctc tat gcc gtg gta ttc gcc att gtg tcg att gcc     192
Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ile Val Ser Ile Ala
    50                  55                  60

ctg att tac ttc ggc agt aca gga atc tgg ccg ctc cag tgg att ggt     240
Leu Ile Tyr Phe Gly Ser Thr Gly Ile Trp Pro Leu Gln Trp Ile Gly
65                  70                  75                  80

gca ggc atg acc gct tat ggt tta ctg tat ttt atg gtc cac gac gga     288
Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
                85                  90                  95

ctg gta cac cag cgc tgg ccg ttc cgc tac ata ccg cgc aaa ggc tac     336
Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
                100                 105                 110

ctg aaa cgg tta tac atg gcc cac cgt atg cat cat gct gta agg gga     384
Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
            115                 120                 125

aaa gag ggc tgc gtg tcc ttt ggt ttt ctg tac gcg cca ccg tta tct     432
Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
        130                 135                 140

aaa ctt cag gcg acg ctg aga gaa agg cat gcg gct aga tcg ggc gct     480
Lys Leu Gln Ala Thr Leu Arg Glu Arg His Ala Ala Arg Ser Gly Ala
145                 150                 155                 160

gcc aga gat gag cag gac ggg gtg gat acg tct tca tcc ggg aag taa     528
Ala Arg Asp Glu Gln Asp Gly Val Asp Thr Ser Ser Ser Gly Lys
                165                 170                 175


<210>   12
<211>   175
<212>   PRT
```

<213> Pantoea stewartii

<400> 12

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Val Gly
1               5                   10                  15

Met Glu Val Val Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
                20                  25                  30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
        35                  40                  45

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ile Val Ser Ile Ala
    50                  55                  60

Leu Ile Tyr Phe Gly Ser Thr Gly Ile Trp Pro Leu Gln Trp Ile Gly
65                  70                  75                  80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
                85                  90                  95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
                100                 105                 110

Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
                115                 120                 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
    130                 135                 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Ala Ala Arg Ser Gly Ala
145                 150                 155                 160

Ala Arg Asp Glu Gln Asp Gly Val Asp Thr Ser Ser Ser Gly Lys
                165                 170                 175

<210>   13
<211>   25
<212>   DNA
<213>   Artificial sequence

<220>
<223>   Primer used to amplify crt gene cluster.

<400>   13
atgacggtct gcgcaaaaaa acacg                                              25

<210>   14
<211>   28
<212>   DNA
<213>   Artificial sequence

<220>
<223>   Primer used to amplify crt gene cluster.

<400> 14
gagaaattat gttgtggatt tggaatgc                                           28


<210>  15
<211>  21
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer Tn5PCRF

<400>  15
gctgagttga aggatcagat c                                                  21


<210>  16
<211>  21
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer Tn5PCRR

<400>  16
cgagcaagac gtttcccgtt g                                                  21


<210>  17
<211>  25
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer Kan-2 FP-1

<400>  17
acctacaaca aagctctcat caacc                                              25


<210>  18
<211>  25
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer Kan-2 RP-1

<400>  18
gcaatgtaac atcagagatt ttgag                                              25


<210>  19
<211>  3159
<212>  DNA
<213>  Escherichia coli

<400>  19
atgcctgtta taactcttcc tgatggcagc caacgccatt acgatcacgc tgtaagcccc      60

atggatgttg cgctggacat tggtccaggt ctggcgaaag cctgtatcgc agggcgcgtt     120

aatggcgaac tggttgatgc ttgcgatctg attgaaaacg acgcacaact gtcgatcatt     180

accgccaaag acgaagaagg tctggagatc attcgtcact cctgtgcgca cctgttaggg     240

cacgcgatta aacaactttg gccgcatacc aaaatggcaa tcggcccggt tattgacaac     300

```
ggtttttatt acgacgttga tcttgaccgc acgttaaccc aggaagatgt cgaagcactc    360

gagaagcgga tgcatgagct tgctgagaaa aactacgacg tcattaagaa gaaagtcagc    420

tggcacgaag cgcgtgaaac tttcgccaac cgtggggaga gctacaaagt ctccattctt    480

gacgaaaaca tcgcccatga tgacaagcca ggtctgtact tccatgaaga atatgtcgat    540

atgtgccgcg gtccgcacgt accgaacatg cgtttctgcc atcatttcaa actaatgaaa    600

acggcagggg cttactggcg tggcgacagc aacaacaaaa tgttgcaacg tatttacggt    660

acggcgtggg cagacaaaaa agcacttaac gcttacctgc agcgcctgga agaagccgcg    720

aaacgcgacc accgtaaaat cggtaaacag ctcgacctgt accatatgca ggaagaagcg    780

ccgggtatgg tattctggca caacgacggc tggaccatct tccgtgaact ggaagtgttt    840

gttcgttcta aactgaaaga gtaccagtat caggaagtta aaggtccgtt catgatggac    900

cgtgtcctgt gggaaaaaac cggtcactgg acaactaca aagatgcaat gttcaccaca    960

tcttctgaga accgtgaata ctgcattaag ccgatgaact gcccgggtca cgtacaaatt   1020

ttcaaccagg ggctgaagtc ttatcgcgat ctgccgctgc gtatggccga gtttggtagc   1080

tgccaccgta acgagccgtc aggttcgctg catggcctga tgcgcgtgcg tggatttacc   1140

caggatgacg cgcatatctt ctgtactgaa gaacaaattc gcgatgaagt taacggatgt   1200

atccgtttag tctatgatat gtacagcact tttggcttcg agaagatcgt cgtcaaactc   1260

tccactcgtc ctgaaaaacg tattggcagc gacgaaatgt gggatcgtgc tgaggcggac   1320

ctggcggttg cgctggaaga aaacaacatc ccgtttgaat atcaactggg tgaaggcgct   1380

ttctacggtc cgaaaattga atttaccctg tatgactgcc tcgatcgtgc atggcagtgc   1440

ggtacagtac agctggactt ctctttgccg tctcgtctga gcgcttctta tgtaggcgaa   1500

gacaatgaac gtaaagtacc ggtaatgatt caccgcgcaa ttctggggtc gatggaacgt   1560

ttcatcggta tcctgaccga agagttcgct ggtttcttcc cgacctggct tgcgccggtt   1620

caggttgtta tcatgaatat taccgattca cagtctgaat acgttaacga attgacgcaa   1680

aaactatcaa atgcgggcat tcgtgttaaa gcagacttga gaaatgagaa gattggcttt   1740

aaaatccgcg agcacacttt gcgtcgcgtc ccatatatgc tggtctgtgg tgataaagag   1800

gtggaatcag gcaaagttgc cgttcgcacc cgccgtggta aagacctggg aagcatggac   1860

gtaaatgaag tgatcgagaa gctgcaacaa gagattcgca ccgcagtct taaacctgtc   1920

tcttatacac atctcaacca tcatcgatga attgtgtctc aaaatctctg atgttacatt   1980

gcacaagata aaaatatatc atcatgaaca ataaaactgt ctgcttacat aaacagtaat   2040

acaaggggtg ttatgagcca tattcaacgg gaaacgtctt gctcgaggcc gcgattaaat   2100

tccaacatgg atgctgattt atatgggtat aaatgggctc gcgataatgt cgggcaatca   2160

ggtgcgacaa tctatcgatt gtatgggaag cccgatgcgc cagagttgtt tctgaaacat   2220

ggcaaaggta gcgttgccaa tgatgttaca gatgagatgg tcagactaaa ctggctgacg   2280

gaatttatgc ctcttccgac catcaagcat tttatccgta ctcctgatga tgcatggtta   2340
```

```
ctcaccactg cgatccccgg aaaaacagca ttccaggtat tagaagaata tcctgattca   2400

ggtgaaaata ttgttgatgc gctggcagtg ttcctgcgcc ggttgcattc gattcctgtt   2460

tgtaattgtc cttttaacag cgatcgcgta tttcgtctcg ctcaggcgca atcacgaatg   2520

aataacggtt tggttgatgc gagtgatttt gatgacgagc gtaatggctg gcctgttgaa   2580

caagtctgga aagaaatgca taaacttttg ccattctcac cggattcagt cgtcactcat   2640

ggtgatttct cacttgataa ccttattttt gacgagggga aattaatagg ttgtattgat   2700

gttggacgag tcggaatcgc agaccgatac caggatcttg ccatcctatg gaactgcctc   2760

ggtgagtttt ctccttcatt acagaaacgg cttttcaaa aatatggtat tgataatcct   2820

gatatgaata aattgcagtt tcatttgatg ctcgatgagt ttttctaatc agaattggtt   2880

aattggttgt aacactggca gagcattacg ctgacttgac gggacggcgg ctttgttgaa   2940

taaatcgaac ttttgctgag ttgaaggatc agatcacgca tcttcccgac aacgcagacc   3000

gttccgtggc aaagcaaaag ttcaaaatca ccaactggtc cacctacaac aaagctctca   3060

tcaaccgtgg cggggatcct ctagagtcga cctgcaggca tgcaagcttc agggttgaga   3120

tgtgtataag agacaggtct taaacaattg gaggaataa                          3159


<210>   20
<211>   3171
<212>   DNA
<213>   Escherichia coli

<400>   20
atgatgagtt atgtagactg gccgccatta attttgaggc acacgtacta catggctgaa     60

ttcgaaacca cttttgcaga tctgggcctg aaggctccta tccttgaagc ccttaacgat    120

ctgggttacg aaaaaccatc tccaattcag gcagagtgta ttccacatct gctgaatggc    180

cgcgacgttc tgggtatggc ccagacgggg agcggaaaaa ctgcagcatt ctctttacct    240

ctgttgcaga tcttgatcc tgagctgaaa gcaccacaga ttctggtgct ggcaccgacc    300

cgcgaactgg cggtacaggt tgctgaagca atgacggatt ctctaaaca catgcgcggc    360

gtaaatgtgg ttgctctgta cggcggccag cgttatgacg tgcaattacg cgccctgcgt    420

caggggccgc agatcgttgt cggtactccg gccgtctgc tggaccacct gaaacgtggc    480

actctggacc tctctaaact gagcggtctg gttctggatg aagctgacga atgctgcgc    540

atgggcttca tcgaagacgt tgaaaccatt atggcgcaga tcccggaagg tcatcagacc    600

gctctgttct ctgcaaccat gccggaagcg attcgtcgca ttacccgccg ctttatgaaa    660

gagccgcagg aagtgcgcat tcagtccagc gtgactaccc gtcctgacat cagccagagc    720

tactggactg tctggggtat gcgcaaaaac gaagcactgg tacgctgtct cttatacaca    780

tctcaaccat catcgatgaa ttgtgtctca aaatctctga tgttacattg cacaagataa    840

aaatatatca tcatgaacaa taaaactgtc tgcttacata aacagtaata caaggggtgt    900

tatgagccat attcaacggg aaacgtcttg ctcgaggccg cgattaaatt ccaacatgga    960

tgctgattta tatgggtata aatgggctcg cgataatgtc gggcaatcag gtgcgacaat   1020
```

ctatcgattg tatgggaagc ccgatgcgcc agagttgttt ctgaaacatg gcaaaggtag     1080

cgttgccaat gatgttacag atgagatggt cagactaaac tggctgacgg aatttatgcc     1140

tcttccgacc atcaagcatt ttatccgtac tcctgatgat gcatggttac tcaccactgc     1200

gatccccgga aaaacagcat tccaggtatt agaagaatat cctgattcag gtgaaaatat     1260

tgttgatgcg ctggcagtgt tcctgcgccg gttgcattcg attcctgttt gtaattgtcc     1320

ttttaacagc gatcgcgtat ttcgtctcgc tcaggcgcaa tcacgaatga ataacggttt     1380

ggttgatgcg agtgattttg atgacgagcg taatggctgg cctgttgaac aagtctggaa     1440

agaaatgcat aaacttttgc cattctcacc ggattcagtc gtcactcatg gtgatttctc     1500

acttgataac cttattttg acgaggggaa attaataggt tgtattgatg ttggacgagt     1560

cggaatcgca gaccgatacc aggatcttgc catcctatgg aactgcctcg gtgagttttc     1620

tccttcatta cagaaacggc tttttcaaaa atatggtatt gataatcctg atatgaataa     1680

attgcagttt catttgatgc tcgatgagtt tttctaatca gaattggtta attggttgta     1740

acactggcag agcattacgc tgacttgacg ggacggcggc tttgttgaat aaatcgaact     1800

tttgctgagt tgaaggatca gatcacgcat cttcccgaca acgcagaccg ttccgtggca     1860

aagcaaaagt tcaaaatcac caactggtcc acctacaaca aagctctcat caaccgtggc     1920

ggggatcctc tagagtcgac ctgcaggcat gcaagcttca gggttgagat gtgtataaga     1980

gacagactgg tacgtttcct ggaagcggaa gattttgatg cggcgattat cttcgttcgt     2040

accaaaaacg cgactctgga agtggctgaa gctcttgagc gtaacggcta caacagcgcc     2100

gcgctgaacg gtgacatgaa ccaggcgctg cgtgaacaga cactggaacg cctgaaagat     2160

ggtcgtctgg acatcctgat tgcgaccgac gttgcagccc gtggcctgga cgttgagcgt     2220

atcagcctgg tagttaacta cgatatcccg atggattctg agtcttacgt tcaccgtatc     2280

ggtcgtaccg tcgtgcggg tcgtgctggc cgcgcgctgc tgttcgttga gaaccgcgag     2340

cgtcgtctgc tgcgcaacat tgaacgtact atgaagctga ctattccgga agtagaactg     2400

ccgaacgcag aactgctagg caaacgccgt ctggaaaaat cgccgctaa agtacagcag     2460

cagctggaaa gcagcgatct ggatcaatac cgcgcactgc tgagcaaaat tcagccgact     2520

gctgaaggtg aagagctgga tctcgaaact ctggctgcgg cactgctgaa aatggcacag     2580

ggtgaacgta ctctgatcgt accgccagat gcgccgatgc gtccgaaacg tgaattccgt     2640

gaccgtgatg accgtggtcc gcgcgatcgt aacgaccgtg cccgcgtgg tgaccgtgaa     2700

gatcgtccgc gtcgtgaacg tcgtgatgtt ggcgatatgc agctgtaccg cattgaagtg     2760

ggccgcgatg atggtgttga agttcgtcat atcgttggtg cgattgctaa cgaaggcgac     2820

atcagcagcc gttacattgg taacatcaag ctgtttgctt ctcactccac catcgaactg     2880

ccgaaaggta tgccgggtga agtgctgcaa cactttacgc gcactcgcat tctcaacaag     2940

ccgatgaaca tgcagttact gggcgatgca cagccgcata ctggcggtga gcgtcgtggc     3000

ggtggtcgtg gtttcggtgg cgaacgtcgt gaaggcggtc gtaacttcag cggtgaacgc     3060

```
cgtgaaggtg gccgtggtga tggtcgtcgt tttagcggcg aacgtcgtga aggccgcgct    3120

ccgcgtcgtg atgattctac cggtcgtcgt cgtttcggtg gtgatgcgta a              3171
```

```
<210>    21
<211>    2904
<212>    DNA
<213>    Escherichia coli

<400>    21
atgactgaat cttttgctca actctttgaa gagtccttaa aagaaatcga aacccgcccg      60

ggttctatcg ttcgtggcgt tgttgttgct atcgacaaag acgtagtact ggttgacgct     120

ggtctgaaat ctgagtccgc catcccggct gagcagttca aaaacgccca gggcgagctg     180

gaaatccagg taggtgacga agttgacgtt gctctggacg cagtagaaga cggcttcggt     240

gaaactctgc tgtcccgtga aaagctaaa cgtcacgaag cctggatcac gctggaaaaa      300

gcttacgaag atgctgaaac tgttaccggt gttatcaacg gcaaagttaa gggcggcttc     360

actgttgagc tgaacggtat tcgtgcgttc ctgccaggtt ctctggtaga cgttcgtccg     420

gtgcgtgaca ctctgcacct ggaaggcaaa gagcttgaat ttaaagtaat caagctggat     480

cagaagcgca acaacgttgt tgtttctcgt cgtgccgtta tcgaatccga aaacagcgca     540

gagcgcgatc agctgctgga aaacctgcag gaaggcatgg aagttaaagg tatcgttaag     600

aacctcactg actacggtgc attcgttgat ctgggcggcg ttgacggcct gctgcacatc     660

actgacatgg cctggaaacg cgttaagcat ccgagcgaaa tcgtcaacgt gggcgacgaa     720

atcactgtta aagtgctgaa gttcgaccgc gaacgtaccc gtgtatccct gggcctgaaa     780

cagctgggcg aagatccgtg ggtagctatc gctaaacgtt atccggaagg taccaaactg     840

actggtcgcg tgaccaacct gaccgactac ggctgcttcg ttgaaatcga agaaggcgtt     900

gaaggcctgg tacacgtttc cgaaatggac tggaccaaca aaaacatcca cccgtccaaa     960

gttgttaacg ttggcgatgt agtggaagtt atggttctgg atatcgacga agaacgtcgt    1020

cgtatctccc tgggtctgaa acagtgcaaa gctaacccgt ggcagcagtt cgcgggaaacc   1080

cacaacaagg cgaccgtgt tgaaggtaaa atcaagtcta tcactgactt cggtatcttc     1140

atcggcttgg acggcggcat cgacggcctg gttcacctgt ctgacatctc ctggaacgtt    1200

gcaggcgaag aagcagttcg tgaatacaaa aaaggcgacg aaatcgctgc agttgttctg    1260

caggttgacg cagaacgtga acgtatctcc ctgggcgtta aacagctcgc agaagatccg    1320

ttcaacaact gggttgctct gaacaagaaa ggcgctatcg taaccggtaa agtaactgca    1380

gttgacgcta aaggcgcaac cgtagaactg gctgacggcg ttgaaggtta cctgcgtgct    1440

tctgaagcat cccgtgaccg cgttgaagac gctaccctgg ttctgagcgt tggcgacgaa    1500

gttgaagcta aattcaccgg cgttgatcgt aaaaaccgcg caatcagcct gtctgttcgt    1560

gcgaaagacg aagctgacga gaaagatgca atcgcaactg tctcttatac acatctcaac    1620

cctgaagctt gcatgcctgc aggtcgactc tagaggatcc ccgccacggt tgatgagagc    1680
```

```
tttgttgtag gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc   1740

gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca   1800

aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat   1860

tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta   1920

tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag   1980

ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata   2040

caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg   2100

acgactgaat ccggtgagaa tggcaaaagt ttatgcattt ctttccagac ttgttcaaca   2160

ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt   2220

gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga   2280

atcgaatgca accggcgcag gaacactgcc agcgcatcaa caatattttc acctgaatca   2340

ggatattctt ctaatacctg gaatgctgtt tttccgggga tcgcagtggt gagtaaccat   2400

gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc   2460

cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctacctttt gccatgtttc   2520

agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc   2580

ccgacattat cgcgagccca tttatadcca tataaatcag catccatgtt ggaatttaat   2640

cgcggcctcg agcaagacgt ttcccgttga atatggctca taacacccct tgtattactg   2700

tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa   2760

catcagagat tttgagacac aattcatcga tgatggttga gatgtgtata agagacagca   2820

atcgcaactg ttaacaaaca ggaagatgca aacttctcca caacgcaat ggctgaagct   2880

ttcaaagcag ctaaaggcga gtaa                                         2904
```

```
<210>   22
<211>   5454
<212>   DNA
<213>   Escherichia coli

<400>   22
gtgaaagatt tattaaagtt tctgaaagcg cagactaaaa ccgaagagtt tgatgcgatc     60

aaaattgctc tggcttcgcc agacatgatc cgttcatggt ctttcggtga agttaaaaag    120

ccggaaacca tcaactaccg tacgttcaaa ccagaacgtg acggcctttt ctgcgcccgt    180

atctttgggc cggtaaaaga ttacgagtgc ctgtgcggta agtacaagcg cctgaaacac    240

cgtggcgtca tctgtgagaa gtgcggcgtt gaagtgaccc agactaaagt acgccgtgag    300

cgtatgggcc acatcgaact ggcttccccg actgcgcaca tctggttcct gaaatcgctg    360

ccgtcccgta tcggtctgct gctcgatatg ccgctgcgcg atatcgaacg cgtactgtac    420

tttgaatcct atgtggttat cgaaggcggt atgaccaacc tggaacgtca gcagatcctg    480

actgaagagc agtatctgga cgcgctggaa gagttcggtg acgaattcga cgcgaagatg    540

ggggcggaag caatccaggc tctgctgaag agcatggatc tggagcaaga gtgcgaacag    600
```

```
ctgcgtgaag agctgaacga aaccaactcc gaaaccaagc gtaaaaagct gaccaagcgt    660

atcaaactgc tggaagcgtt cgttcagtct ggtaacaaac cagagtggat gatcctgacc    720

gttctgccgg tactgccgcc agatctgcgt ccgctggttc cgctggatgg tggtcgtttc    780

gcgacttctg acctgaacga tctgtatcgt cgcgtcatta accgtaacaa ccgtctgaaa    840

cgtctgctgg atctggctgc gccggacatc atcgtacgta acgaaaaacg tatgctgcag    900

gaagcggtag acgccctgct ggataacggt cgtcgcggtc gtgcgatcac cggttctaac    960

aagcgtcctc tgaaatcttt ggccgacatg atcaaaggta aacagggtcg tttccgtcag    1020

aacctgctcg gtaagcgtgt tgactactcc ggtcgttctg taatcaccgt aggtccatac    1080

ctgcgtctgc atcagtgcgg tctgccgaag aaaatggcac tggagctgtt caaaccgttc    1140

atctacggca agctggaact gcgtggtctt gctaccacca ttaaagctgc gaagaaaatg    1200

gttgagcgcg aagaagctgt cgtttgggat atcctggacg aagttatccg cgaacacccg    1260

gtactgctga accgtgcacc gactctgcac cgtctgggta tccaggcatt tgaaccggta    1320

ctgatcgaag gtaaagctat ccagctgcac ccgctggttt gtgcggcata taacgccgac    1380

ttcgatggtg accagatggc tgttcacgta ccgctgacgc tggaagccca gctggaagcg    1440

cgtgcgctga tgatgtctac caacaacatc ctgtccccgg cgaacggcga accaatcatc    1500

gttccgtctc aggacgttgt actgggtctg tactacatga cccgtgactg tgttaacgcc    1560

aaaggcgaag gcatggtgct gactggcccg aaagaagcag aacgtctgta tcgctctggt    1620

ctggcttctc tgcatgcgcg cgttaaagtg cgtatcaccg agtatgaaaa agatgctaac    1680

ggtgaattag tagcgaaaac cagcctgaaa gacacgactg ttggccgtgc cattctgtgg    1740

atgattgtac cgaaaggtct gccttactcc atcgtcaacc aggcgctggg taaaaaagca    1800

atctccaaaa tgctgaacac ctgctaccgc attctcggtc tgaaaccgac cgttattttt    1860

gcggaccaga tcatgtacac cggcttcgcc tatgcagcgc gttctggtgc atctgttggt    1920

atcgatgaca tggtcatccc ggagaagaaa cacgaaatca tctccgaggc agaagcagaa    1980

gttgctgaaa ttcaggagca gttccagtct ggtctggtaa ctgcgggcga acgctacaac    2040

aaagttatcg atatctgggc tgcggcgaac gatcgtgtat ccaaagcgat gatggataac    2100

ctgcaaactg aaaccgtgat taaccgtgac ggtcaggaag agaagcaggt ttccttcaac    2160

agcatctaca tgatggccga ctccggtgcg cgtggttctg cggcacagat tcgtcagctt    2220

gctggtatgc gtggtctgat ggcgaagccg gatggctcca tcatcgaaac gccaatcacc    2280

gcgaacttcc gtgaaggtct gaacgtactc cagtacttca tctccaccca cggtgctcgt    2340

aaaggtctgg cggataccgc actgaaaact gcgaactccg gttacctgac tcgtcgtctg    2400

gttgacgtgg cgcaggacct ggtggttacc gaagacgatt gtggtaccca tgaaggtatc    2460

atgatgactc cggttatcga gggtggtgac gttaaagagc gctgcgcga tcgcgtactg    2520

ggtcgtgtaa ctgctgaaga cgttctgaag ccgggtactg ctgatatcct cgttccgcgc    2580

aacacgctgc tgcacgaaca gtggtgtgac ctgctggaag agaactctgt cgacgcggtt    2640
```
Page 24

```
aaagtacgtt ctgttgtatc ttgtgacacc gactttggtg tatgtgcgca ctgctacggt     2700

cgtgacctgg cgcgtggcca catcatcaac aagggtgaag caatcggtgt tatcgcggca     2760

cagtccatcg gtgaaccggg tacacagctg accatgcgta cgttccacat cggtggtgcg     2820

gcatctcgtg cggctgctga atccagcatc caagtgaaaa acaaaggtag catcaagctc     2880

agcaacgtga agtcggttgt gaactccagc ggtaaactgg ttatcacttc ccgtaatact     2940

gaactgaaac tgatcgacga attcggtcgt actaaagaaa gctacaaagt accttacggt     3000

gcggtactgg cgaaaggcga tggcgaacag gttgctggcg gcgaaaccgt tgcaaactgg     3060

gacccgcaca ccatgccggt tatcaccgaa gtaagcggtt ttgtacgctt tactgacatg     3120

atcgacggcc agaccattac gcgtcagacc gacgaactga ccggtctgtc ttcgctggtg     3180

gttctggatt ccgcagaacg taccgcaggt ggtaaagatc tgcgtccggc actgaaaatc     3240

gttgatgctc agggtaacga cgttctgatc ccaggtaccg atatgccagc gcagtacttc     3300

ctgccgggta aagcgattgt tcagctggaa gatggcgtac agatcagctc tggtgacacc     3360

ctggcgcgta ttccgcagga atccggcggt accaaggaca tcaccggtgg tctgccgcgc     3420

gttgcggacc tgttcgaagc acgtcgtccg aaagagccgg caatcctggc tgaaatcagc     3480

ggtatcgttt ccttcggtaa agaaaccaaa ggtaaacgtc gtctggttat caccccggta     3540

gacggtagcg atccgtacga agagatgatt ccgaaatggc gtcagctcaa cgtgttcgaa     3600

ggtgaacgtg tagaacgtgg tgacgtaatt tccgacggtc cggaagcgcc gcacgacatt     3660

ctgcgtctgc gtggtgttca tgctgttact cgttacatcg ttaacgaagt acaggacgta     3720

taccgtctgc agggcgttaa gattaacgat aaacacatcg aagttatcgt tcgtcagatg     3780

ctgcgtaaag ctaccatcgt taacgcgggt agctccgact tcctggaagg cgaacaggtt     3840

gaatactctc gcgtcaagat cgcaaaccgc gaactggaag cgaacggcaa agtgggtgca     3900

acttactccc gcgatctgct gggtatcacc aaagcgtctc tggcaaccga gtccttcatc     3960

tccgcggcat cgttccagga gaccactcgc gtgctgaccg aagcagccgt tgcgggcaaa     4020

cgcgacgaac tgcgcggcct gaaagagaac gttatcgtgg gtcgtctgat cccggcaggt     4080

accggttacg cgtaccacca ggatcgtatg cgtcgccgtg ctgcgggtga agctctgtct     4140

cttatacaca tctcaaccct gaagcttgca tgcctgcagg tcgactctag aggatccccg     4200

ccacggttga tgagagcttt gttgtaggtg gaccagttgg tgattttgaa cttttgcttt     4260

gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa ctcagcaaaa     4320

gttcgattta ttcaacaaag ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt     4380

acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt     4440

tattcatatc aggattatca ataccatatt tttgaaaaag ccgtttctgt aatgaaggag     4500

aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcggtct gcgattccga     4560

ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagtg     4620

agaaatcacc atgagtgacg actgaatccg gtgagaatgg caaaagttta tgcatttctt     4680
```

```
tccagacttg ttcaacaggc cagccattac gctcgtcatc aaaatcactc gcatcaacca    4740

aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag    4800

gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa    4860

tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttt ccggggatcg    4920

cagtggtgag taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag    4980

gcataaattc cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc    5040

tacctttgcc atgtttcaga aacaactctg gcgcatcggg cttcccatac aatcgataga    5100

ttgtcgcacc tgattgcccg acattatcgc gagcccattt atacccatat aaatcagcat    5160

ccatgttgga atttaatcgc ggcctcgagc aagacgtttc ccgttgaata tggctcataa    5220

caccccttgt attactgttt atgtaagcag acagttttat tgttcatgat gatatatttt    5280

tatcttgtgc aatgtaacat cagagatttt gagacacaat tcatcgatga tggttgagat    5340

gtgtataaga dacagggtga agctccggct gcaccgcagg tgactgcaga agacgcatct    5400

gccagcctgg cagaactgct gaacgcaggt ctgggcggtt ctgataacga gtaa         5454
```

<210>    23
<211>    1845
<212>    DNA
<213>    Escherichia coli

<400>    23
```
atgggcaaaa catctatgat acacgcaatt gtggatcaat atagtcactg tgaatgggtg      60

gaaaatagca tgagtgccaa tgaaaacaac ctgatttgga tcgatcttga gatgaccggt     120

ctggatcccg agcgcgatcg cattattgag attgccacgc tggtgaccga tgccaacctg     180

aatattctgg cagaagggcc gaccattgca gtacaccagt ctgatgaaca gctggcgctg     240

atggatgact ggaacgtgcg cacccatacc gccagcgggc tggtagagcg cgtgaaagcg     300

agcacgatgg gcgatcggga agctgaactg gcaacgctcg aatttttaaa acagtgggtg     360

cctgcgggaa aatcgccgat ttgcggtaac agcatcggtc aggaccgtcg tttcctgttt     420

aaatacatgc cggagctgga agcctacttc cactaccgtt atctcgatgt cagcaccctg     480

aaagagctgg cgcgccgctg gaagccggaa attctggatg gtttttaccaa gcaggggacg     540

catcaggcga tggatgatat ccgtgaatcg gtggcggagc tggcttacta cctgtctctt     600

atacacatct caaccctgaa gcttgcatgc ctgcaggtcg actctagagg atccccgcca     660

cggttgatga gagctttgtt gtaggtggac cagttggtga ttttgaactt ttgctttgcc     720

acggaacggt ctgcgttgtc gggaagatgc gtgatctgat ccttcaactc agcaaaagtt     780

cgatttattc aacaaagccg ccgtcccgtc aagtcagcgt aatgctctgc cagtgttaca     840

accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac tgcaatttat     900

tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat gaaggagaaa     960

actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg attccgactc    1020
```

```
gtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta tcaagtgaga   1080

aatcaccatg agtgacgact gaatccggtg agaatggcaa aagtttatgc atttctttcc   1140

agacttgttc aacaggccag ccattacgct cgtcatcaaa atcactcgca tcaaccaaac   1200

cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac   1260

aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat   1320

tttcacctga atcaggatat tcttctaata cctggaatgc tgtttttccg gggatcgcag   1380

tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc ggaagaggca   1440

taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac   1500

ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat cgatagattg   1560

tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca   1620

tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg ctcataacac   1680

cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat atattttat   1740

cttgtgcaat gtaacatcag agattttgag acacaattca tcgatgatgg ttgagatgtg   1800

tataagagac aggcttacta ccgcgagcat tttatcaagc tgtaa               1845
```

```
<210>  24
<211>  2334
<212>  DNA
<213>  Escherichia coli

<400>  24
atgaagccaa tttttagccg tggcccgtcg ctacagattc gccttattct ggcggtgctg    60

gtggcgctcg gcattattat tgccgacagc cgcctgggga cgttcagtca aatccgtact   120

tatatggata ccgccgtcag tcctttctac tttgtttcca atgctcctcg tgaattgctg   180

gatggcgtat cgcagacgct ggcctcgcgt gaccaattag aacttgaaaa ccgggcgtta   240

cgtcaggaac tgttgctgaa aaacagtgaa ctgctgatgc ttggacaata caaacaggag   300

aacgcgcgtc tgcgcgagct gctgggttcc ccgctgcgtc aggatgagca gaaaatggtg   360

actcaggtta tctccacggt taacgatcct tatagcgatc aagttgttat cgataaaggt   420

agcgttaatg gcgtttatga aggccagccg gtcatcagcg acaaaggtgt tgttggtcag   480

gtggtggccg tcgctaaact gaccagtcgc gtgctgctga tttgtgatgc gacccacgcg   540

ctgccaatcc aggtgctgcg caacgatatc gcgtaattg cagccggtaa cggttgtacg   600

gatgatttgc agcttgagca tctgccggcg aatacggata ttcgtgttgg tgatgtgctg   660

gtgacttccg gtctgggcgg tcgtttcccg gaaggctatc cggtcgcggt tgtctcttcc   720

gtaaaactcg atacccagcg cgcttatact gtgattcagg cgcgtccgac tgcagggctg   780

caacgtttgc gttatctgct gctgctgtgg ggggcagatc gtaacggcgc taacccgatg   840

acgccggaag aggtgcatcg tgttgctaat gaacgtctga tgcagatgat gccgcaggta   900

ttgccttcgc cagacgcgat ggggccaaag ttacctgaac cggcaacggg gatcgctcag   960

ccgactccgc agcaaccggc gacaggaaat gcagctactg cgcctgctgc gccgacacag   1020
```

```
cctctgtctc ttatacacat ctcaaccatc atcgatgaat tgtgtctcaa aatctctgat    1080

gttacattgc acaagataaa aatatatcat catgaacaat aaaactgtct gcttacataa    1140

acagtaatac aaggggtgtt atgagccata ttcaacggga aacgtcttgc tcgaggccgc    1200

gattaaattc caacatggat gctgatttat atgggtataa atgggctcgc gataatgtcg    1260

ggcaatcagg tgcgacaatc tatcgattgt atgggaagcc cgatgcgcca gagttgtttc    1320

tgaaacatgg caaaggtagc gttgccaatg atgttacaga tgagatggtc agactaaact    1380

ggctgacgga atttatgcct cttccgacca tcaagcattt tatccgtact cctgatgatg    1440

catggttact caccactgcg atccccggaa aaacagcatt ccaggtatta gaagaatatc    1500

ctgattcagg tgaaaatatt gttgatgcgc tggcagtgtt cctgcgccgg ttgcattcga    1560

ttcctgtttg taattgtcct tttaacagcg atcgcgtatt tcgtctcgct caggcgcaat    1620

cacgaatgaa taacggtttg gttgatgcga gtgattttga tgacgagcgt aatggctggc    1680

ctgttgaaca agtctggaaa gaaatgcata aacttttgcc attctcaccg gattcagtcg    1740

tcactcatgg tgatttctca cttgataacc ttatttttga cgaggggaaa ttaataggtt    1800

gtattgatgt tggacgagtc ggaatcgcag accgatacca ggatcttgcc atcctatgga    1860

actgcctcgg tgagttttct ccttcattac agaaacggct ttttcaaaaa tatggtattg    1920

ataatcctga tatgaataaa ttgcagtttc atttgatgct cgatgagttt ttctaatcag    1980

aattggttaa ttggttgtaa cactggcaga gcattacgct gacttgacgg gacggcggct    2040

ttgttgaata aatcgaactt ttgctgagtt gaaggatcag atcacgcatc ttcccgacaa    2100

cgcagaccgt tccgtggcaa agcaaaagtt caaaatcacc aactggtcca cctacaacaa    2160

agctctcatc aaccgtggcg gggatcctct agagtcgacc tgcaggcatg caagcttcag    2220

ggttgagatg tgtataagag acagacacag cctgctgcta atcgctctcc acaaagggct    2280

acgccgccgc aaagtggtgc tcaaccgcct gcgcgtgcgc cgggagggca atag          2334
```

<210> 25
<211> 2676
<212> DNA
<213> Escherichia coli

<400> 25

```
atgcgaagtg aacagatttc tggctcgtca ctcaatccgt cttgtcgttt cagttcctgt      60

ctcttataca catctcaacc atcatcgatg aattgtgtct caaaatctct gatgttacat     120

tgcacaagat aaaaatatat catcatgaac aataaaactg tctgcttaca taaacagtaa     180

tacaaggggt gttatgagcc atattcaacg ggaaacgtct tgctcgaggc cgcgattaaa     240

ttccaacatg gatgctgatt tatatgggta aatgggctcg cgataatgtc gggcaatc       300

aggtgcgaca atctatcgat tgtatgggaa gcccgatgcg ccagagttgt ttctgaaaca     360

tggcaaaggt agcgttgcca atgatgttac agatgagatg gtcagactaa actggctgac     420

ggaatttatg cctcttccga ccatcaagca ttttatccgt actcctgatg atgcatggtt     480
```

```
actcaccact gcgatccccg gaaaaacagc attccaggta ttagaagaat atcctgattc   540

aggtgaaaat attgttgatg cgctggcagt gttcctgcgc cggttgcatt cgattcctgt   600

ttgtaattgt cctttaaca gcgatcgcgt atttcgtctc gctcaggcgc aatcacgaat   660

gaataacggt ttggttgatg cgagtgattt tgatgacgag cgtaatggct ggcctgttga   720

acaagtctgg aaagaaatgc ataaactttt gccattctca ccggattcag tcgtcactca   780

tggtgatttc tcacttgata accttatttt tgacgagggg aaattaatag gttgtattga   840

tgttggacga gtcggaatcg cagaccgata ccaggatctt gccatcctat ggaactgcct   900

cggtgagttt tctccttcat tacagaaacg cttttttcaa aaatatggta ttgataatcc   960

tgatatgaat aaattgcagt ttcatttgat gctcgatgag tttttctaat cagaattggt   1020

taattggttg taacactggc agagcattac gctgacttga cgggacggcg ctttgttga   1080

ataaatcgaa cttttgctga gttgaaggat cagatcacgc atcttcccga caacgcagac   1140

cgttccgtgg caaagcaaaa gttcaaaatc accaactggt ccacctacaa caaagctctc   1200

atcaaccgtg gcggggatcc tctagagtcg acctgcaggc atgcaagctt cagggttgag   1260

atgtgtataa gagacagttt cagttctgcg tactctcctg tgaccaggca gcgaaaagac   1320

atgagtcgat gaccgtaaac aggcatggat gatcctgcca taccattcac aacattaagt   1380

tcgagattta ccccaagttt aagaactcac accactatga atcttaccga attaaagaat   1440

acgccggttt ctgagctgat cactctcggc gaaaatatgg ggctggaaaa cctggctcgt   1500

atgcgtaagc aggacattat ttttgccatc ctgaagcagc acgcaaagag tggcgaagat   1560

atctttggtg atggcgtact ggagatattg caggatggat ttggtttcct ccgttccgca   1620

gacagctcct acctcgccgg tcctgatgac atctacgttt cccctagcca aatccgccgt   1680

ttcaacctcc gcactggtga taccatctct ggtaagattc gcccgccgaa agaaggtgaa   1740

cgctattttg cgctgctgaa agttaacgaa gttaacttcg acaaacctga aaacgcccgc   1800

aacaaaatcc tctttgagaa cttaaccccg ctgcacgcaa actctcgtct gcgtatggaa   1860

cgtggtaacg gttctactga agatttaact gctcgcgtac tggatctggc atcacctatc   1920

ggtcgtggtc agcgtggtct gattgtggca ccgccgaaag ccggtaaaac catgctgctg   1980

cagaacattg ctcagagcat tgcttacaac cacccggatt gtgtgctgat ggttctgctg   2040

atcgacgaac gtccggaaga agtaaccgag atgcagcgtc tggtaaaagg tgaagttgtt   2100

gcttctacct ttgacgaacc cgcatctcgc cacgttcagg ttgcggaaat ggtgatcgag   2160

aaggccaaac gcctggttga gcacaagaaa gacgttatca ttctgctcga ctccatcact   2220

cgtctggcgc gcgcttacaa caccgttgtt ccggcgtcag gtaaagtgtt gaccggtggt   2280

gtggatgcca acgccctgca tcgtccgaaa cgcttctttg gtgcggcgcg taacgtggaa   2340

gagggcggca gcctgaccat tatcgcgacg gcgcttatcg ataccggttc taaaatggac   2400

gaagttatct acgaagagtt taaaggtaca ggcaacatgg aactgcacct ctctcgtaag   2460

atcgctgaaa aacgcgtctt cccggctatc gactacaacc gttctggtac ccgtaaagaa   2520
```

```
gagctgctca cgactcagga agaactgcag aaaatgtgga tcctgcgcaa aatcattcac    2580

ccgatgggcg aaatcgatgc aatggaattc ctcattaata aactggcaat gaccaagacc    2640

aatgacgatt tcttcgaaat gatgaaacgc tcataa                             2676


<210>   26
<211>   1746
<212>   DNA
<213>   Escherichia coli

<400>   26
atggattact tcaccctctt tggcttgcct gcccgctatc aactcgatac ccaggcgctg      60

agcctgcgtt ttcaggatct acaacgtcag tatcatcctg ataaattcgc cagcggaagc     120

caggcggaac aactcgccgc cgtacagcaa tctgcaacca ttaaccaggc ctggcaaacg     180

ctgcgtcatc cgttaatgcg cgcggaatat ttgctttctt tgcacggctt tgatctcgcc     240

agcgagcagc atacctgtct cttatacaca tctcaaccat catcgatgaa ttgtgtctca     300

aaatctctga tgttacattg cacaagataa aaatatatca tcatgaacaa taaaactgtc     360

tgcttacata aacagtaata caaggggtgt tatgagccat attcaacggg aaacgtcttg     420

ctcgaggccg cgattaaatt ccaacatgga tgctgattta tatgggtata aatgggctcg     480

cgataatgtc gggcaatcag gtgcgacaat ctatcgattg tatgggaagc ccgatgcgcc     540

agagttgttt ctgaaacatg gcaaaggtag cgttgccaat gatgttacag atgagatggt     600

cagactaaac tggctgacgg aatttatgcc tcttccgacc atcaagcatt ttatccgtac     660

tcctgatgat gcatggttac tcaccactgc gatccccgga aaaacagcat tccaggtatt     720

agaagaatat cctgattcag gtgaaaatat tgttgatgcg ctggcagtgt tcctgcgccg     780

gttgcattcg attcctgttt gtaattgtcc ttttaacagc gatcgcgtat ttcgtctcgc     840

tcaggcgcaa tcacgaatga ataacggttt ggttgatgcg agtgattttg atgacgagcg     900

taatggctgg cctgttgaac aagtctggaa agaaatgcat aaacttttgc cattctcacc     960

ggattcagtc gtcactcatg gtgatttctc acttgataac cttatttttg acgaggggaa    1020

attaataggt tgtattgatg ttggacgagt cggaatcgca gaccgatacc aggatcttgc    1080

catcctatgg aactgcctcg gtgagttttc tccttcatta cagaaacggc tttttcaaaa    1140

atatggtatt gataatcctg atatgaataa attgcagttt catttgatgc tcgatgagtt    1200

tttctaatca gaattggtta attggttgta acactggcag agcattacgc tgacttgacg    1260

ggacggcggc tttgttgaat aaatcgaact tttgctgagt tgaaggatca gatcacgcat    1320

cttcccgaca acgcagaccg ttccgtggca aagcaaaagt tcaaaatcac caactggtcc    1380

acctacaaca aagctctcat caaccgtggc ggggatcctc tagagtcgac ctgcaggcat    1440

gcaagcttca gggttgagat gtgtataaga gacaggcagc atactgtgcg cgacaccgcg    1500

ttcctgatgg aacagttgga gctgcgcgaa gagctggacg agatcgaaca ggcgaaagat    1560

gaagcgcggc tggaaagctt tatcaaacgt gtgaaaaaga tgtttgatac ccgccatcag    1620

ttgatggttg aacagttaga caacgagacg tgggacgcgg cggcggatac cgtgcgtaag    1680
```

```
ctgcgttttc tcgataaact gcgaagcagt gccgaacaac tcgaagaaaa actgctcgat    1740

tttttaa                                                             1746
```

```
<210>  27
<211>  3171
<212>  DNA
<213>  Escherichia coli

<400>  27
atgatgagtt atgtagactg gccgccatta attttgaggc acacgtacta catggctgaa      60

ttcgaaacca cttttgcaga tctgggcctg aaggctccta ccttgaagc ccttaacgat      120

ctgggttacg aaaaaccatc tccaattcag gcagagtgta ttccacatct gctgaatggc      180

cgcgacgttc tgggtatggc ccagacgggg agcggaaaaa ctgcagcatt ctctttacct      240

ctgttgcaga tcttgatcc tgagctgaaa gcaccacaga ttctggtgct ggcaccgacc      300

cgcgaactgg cggtacaggt tgctgaagca atgacggatt tctctaaaca catgcgcggc      360

gtaaatgtgg ttgctctgta cggcggccag cgttatgacg tgcaattacg cgccctgcgt      420

caggggccgc agatcgttgt cggtactccg ggccgtctgc tggaccacct gaaacgtggc      480

actctggacc tctctaaact gagcggtctg gttctggatg aagctgacga aatgctgcgc      540

atgggcttca tcgaagacgt tgaaaccatt atggcgcaga tcccggaagg tcatcagacc      600

gctctgttct ctgcaaccat gccggaagcg attcgtcgca ttacccgccg ctttatgaaa      660

gagccgcagg aagtgcgcat tcagtccagc gtgactaccc gtcctgacat cagccagagc      720

tactggactg tctggggtat gcgcaaaaac gaagcactgg tacgctgtct cttatacaca      780

tctcaaccat catcgatgaa ttgtgtctca aaatctctga tgttacattg cacaagataa      840

aaatatatca tcatgaacaa taaaactgtc tgcttacata aacagtaata caaggggtgt      900

tatgagccat attcaacggg aaacgtcttg ctcgaggccg cgattaaatt ccaacatgga      960

tgctgattta tatgggtata aatgggctcg cgataatgtc gggcaatcag gtgcgacaat     1020

ctatcgattg tatgggaagc ccgatgcgcc agagttgttt ctgaaacatg gcaaaggtag     1080

cgttgccaat gatgttacag atgagatggt cagactaaac tggctgacgg aatttatgcc     1140

tcttccgacc atcaagcatt ttatccgtac tcctgatgat gcatggttac tcaccactgc     1200

gatccccgga aaaacagcat tccaggtatt agaagaatat cctgattcag gtgaaaatat     1260

tgttgatgcg ctggcagtgt tcctgcgccg gttgcattcg attcctgttt gtaattgtcc     1320

ttttaacagc gatcgcgtat ttcgtctcgc tcaggcgcaa tcacgaatga ataacggttt     1380

ggttgatgcg agtgattttg atgacgagcg taatggctgg cctgttgaac aagtctggaa     1440

agaaatgcat aaacttttgc cattctcacc ggattcagtc gtcactcatg gtgatttctc     1500

acttgataac cttattttg acgaggggaa attaataggt tgtattgatg ttggacgagt     1560

cggaatcgca gaccgatacc aggatcttgc catcctatgg aactgcctcg gtgagttttc     1620

tccttcatta cagaaacggc tttttcaaaa atatggtatt gataatcctg atatgaataa     1680
```

```
attgcagttt catttgatgc tcgatgagtt tttctaatca gaattggtta attggttgta    1740

acactggcag agcattacgc tgacttgacg ggacggcggc tttgttgaat aaatcgaact    1800

tttgctgagt tgaaggatca gatcacgcat cttcccgaca acgcagaccg ttccgtggca    1860

aagcaaaagt tcaaaatcac caactggtcc acctacaaca aagctctcat caaccgtggc    1920

ggggatcctc tagagtcgac ctgcaggcat gcaagcttca gggttgagat gtgtataaga    1980

gacagactgg tacgtttcct ggaagcggaa gattttgatg cggcgattat cttcgttcgt    2040

accaaaaacg cgactctgga agtggctgaa gctcttgagc gtaacggcta caacagcgcc    2100

gcgctgaacg gtgacatgaa ccaggcgctg cgtgaacaga cactggaacg cctgaaagat    2160

ggtcgtctgg acatcctgat tgcgaccgac gttgcagccc gtggcctgga cgttgagcgt    2220

atcagcctgg tagttaacta cgatatcccg atggattctg agtcttacgt tcaccgtatc    2280

ggtcgtaccg gtcgtgcggg tcgtgctggc cgcgcgctgc tgttcgttga gaaccgcgag    2340

cgtcgtctgc tgcgcaacat tgaacgtact atgaagctga ctattccgga agtagaactg    2400

ccgaacgcag aactgctagg caaacgccgt ctggaaaaat tcgccgctaa agtacagcag    2460

cagctggaaa gcagcgatct ggatcaatac cgcgcactgc tgagcaaaat tcagccgact    2520

gctgaaggtg aagagctgga tctcgaaact ctggctgcgg cactgctgaa aatggcacag    2580

ggtgaacgta ctctgatcgt accgccagat gcgccgatgc gtccgaaacg tgaattccgt    2640

gaccgtgatg accgtggtcc gcgcgatcgt aacgaccgtg gcccgcgtgg tgaccgtgaa    2700

gatcgtccgc gtcgtgaacg tcgtgatgtt ggcgatatgc agctgtaccg cattgaagtg    2760

ggccgcgatg atggtgttga agttcgtcat atcgttggtg cgattgctaa cgaaggcgac    2820

atcagcagcc gttacattgg taacatcaag ctgtttgctt ctcactccac catcgaactg    2880

ccgaaaggta tgccgggtga agtgctgcaa cactttacgc gcactcgcat tctcaacaag    2940

ccgatgaaca tgcagttact gggcgatgca cagccgcata ctggcggtga gcgtcgtggc    3000

ggtggtcgtg gtttcggtgg cgaacgtcgt gaaggcggtc gtaacttcag cggtgaacgc    3060

cgtgaaggtg gccgtggtga tggtcgtcgt tttagcggcg aacgtcgtga aggccgcgct    3120

ccgcgtcgtg atgattctac cggtcgtcgt cgtttcggtg gtgatgcgta a              3171
```

```
<210>   28
<211>   8609
<212>   DNA
<213>   Artificial sequence

<220>
<223>   Plasmid pPCB15

<400>   28
cgtatggcaa tgaaagacgg tgagctggtg atatgggata gtgttcaccc ttgttacacc     60

gttttccatg agcaaactga aacgtttttca tcgctctgga gtgaataccac cgacgatttc    120

cggcagtttc tacacatata ttcgcaagat gtggcgtgtt acggtgaaaa cctggcctat    180

ttccctaaag ggtttattga gaatatgttt ttcgtctcag ccaatccctg ggtgagtttc    240
```

```
accagttttg atttaaacgt ggccaatatg gacaacttct tcgcccccgt tttcaccatg    300

ggcaaatatt atacgcaagg cgacaaggtg ctgatgccgc tggcgattca ggttcatcat    360

gccgtctgtg atggcttcca tgtcggcaga atgcttaatg aattacaaca gtactgcgat    420

gagtggcagg gcggggcgta atttttttaa ggcagttatt ggtgcctaga aatattttat    480

ctgattaata agatgatctt cttgagatcg ttttggtctg cgcgtaatct cttgctctga    540

aaacgaaaaa accgccttgc agggcggttt ttcgaaggtt ctctgagcta ccaactcttt    600

gaaccgaggt aactggcttg gaggagcgca gtcaccaaaa cttgtccttt cagtttagcc    660

ttaaccggcg catgacttca agactaactc ctctaaatca attaccagtg gctgctgcca    720

gtggtgcttt tgcatgtctt ccgggttgg actcaagacg atagttaccg gataaggcgc    780

agcggtcgga ctgaacgggg ggttcgtgca tacagtccag cttggagcga actgcctacc    840

cggaactgag tgtcaggcgt ggaatgagac aaacgcggcc ataacagcgg aatgacaccg    900

gtaaaccgaa aggcaggaac aggagagcgc acgagggagc cgccagggga aacgcctggt    960

atctttatag tcctgtcggg tttcgccacc actgatttga gcgtcagatt tcgtgatgct   1020

tgtcaggggg gcggagccta tggaaaaacg gctttgccgc ggccctctca cttccctgtt   1080

aagtatcttc ctggcatctt ccaggaaatc tccgccccgt tcgtaagcca tttccgctcg   1140

ccgcagtcga acgaccgagc gtagcgagtc agtgagcgag gaagcggaat atatcctgta   1200

tcacatattc tgctgacgca ccggtgcagc ctttttctc ctgccacatg aagcacttca   1260

ctgacaccct catcagtgcc aacatagtaa gccagtatat acactccgct agcgcccaat   1320

acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc acgacaggtt   1380

tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tcactcatta   1440

ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg   1500

ataacaattt cacacaggaa acagctatga ccatgattac gaattcgagc tcggtaccca   1560

aacgaattcg ccctttttgac ggtctgcgca aaaaaacacg ttcaccttac tggcatttcg   1620

gctgagcagt tgctggctga tatcgatagc cgccttgatc agttactgcc ggttcagggt   1680

gagcgggatt gtgtgggtgc cgcgatgcgt gaaggcacgc tggcaccggg caaacgtatt   1740

cgtccgatgc tgctgttatt aacagcgcgc gatcttggct gtgcgatcag tcacggggga   1800

ttactggatt tagcctgcgc ggttgaaatg gtgcatgctg cctcgctgat tctggatgat   1860

atgccctgca tggacgatgc gcagatgcgt cggggcgtc ccaccattca cacgcagtac   1920

ggtgaacatg tggcgattct ggcggcggtc gctttactca gcaaagcgtt tggggtgatt   1980

gccgaggctg aaggtctgac gccgatagcc aaaactcgcg cggtgtcgga gctgtccact   2040

gcgattggca tgcagggtct ggttcagggc cagtttaagg acctctcgga aggcgataaa   2100

ccccgcagcg ccgatgccat actgctaacc aatcagttta aaaccagcac gctgttttgc   2160

gcgtcaacgc aaatggcgtc cattgcggcc aacgcgtcct gcgaagcgcg tgagaacctg   2220

catcgtttct cgctcgatct cggccaggcc tttcagttgc ttgacgatct taccgatggc   2280
```

```
atgaccgata ccggcaaaga catcaatcag gatgcaggta aatcaacgct ggtcaattta  2340

ttaggctcag gcgcggtcga agaacgcctg cgacagcatt tgcgcctggc cagtgaacac  2400

ctttccgcgg catgccaaaa cggccattcc accacccaac tttttattca ggcctggttt  2460

gacaaaaaac tcgctgccgt cagttaagga tgctgcatga gccattttgc ggtgatcgca  2520

ccgccctttt tcagccatgt tcgcgctctg caaaaccttg ctcaggaatt agtggcccgc  2580

ggtcatcgtg ttacgttttt tcagcaacat gactgcaaag cgctggtaac gggcagcgat  2640

atcggattcc agaccgtcgg actgcaaacg catcctcccg gttccttatc gcacctgctg  2700

cacctggccg cgcacccact cggaccctcg atgttacgac tgatcaatga aatggcacgt  2760

accagcgata tgctttgccg ggaactgccc gccgcttttc atgcgttgca gatagagggc  2820

gtgatcgttg atcaaatgga gccggcaggt gcagtagtcg cagaagcgtc aggtctgccg  2880

tttgtttcgg tggcctgcgc gctgccgctc aaccgcgaac cgggtttgcc tctggcggtg  2940

atgcctttcg agtacggcac cagcgatgcg gctcgggaac gctataccac cagcgaaaaa  3000

atttatgact ggctgatgcg acgtcacgat cgtgtgatcg cgcatcatgc atgcagaatg  3060

ggtttagccc cgcgtgaaaa actgcatcat tgttttctc cactggcaca aatcagccag  3120

ttgatccccg aactggattt tccccgcaaa gcgctgccag actgctttca tgcggttgga  3180

ccgttacggc aaccccaggg gacgccgggg tcatcaactt cttatttttcc gtccccggac  3240

aaaccccgta ttttttgcctc gctgggcacc ctgcagggac atcgttatgg cctgttcagg  3300

accatcgcca aagcctgcga agaggtggat gcgcagttac tgttggcaca ctgtggcggc  3360

ctctcagcca cgcaggcagg tgaactggcc cggggcgggg acattcaggt tgtggatttt  3420

gccgatcaat ccgcagcact ttcacaggca cagttgacaa tcacacatgg tgggatgaat  3480

acggtactgg acgctattgc ttcccgcaca ccgctactgg cgctgccgct ggcatttgat  3540

caacctggcg tggcatcacg aattgtttat catggcatcg gcaagcgtgc gtctcggttt  3600

actaccagcc atgcgctggc gcggcagatt cgatcgctgc tgactaacac cgattacccg  3660

cagcgtatga caaaaattca ggccgcattg cgtctggcag gcggcacacc agccgccgcc  3720

gatattgttg aacaggcgat gcggacctgt cagccagtac tcagtgggca ggattatgca  3780

accgcactat gatctcattc tggtcggtgc cggtctggct aatggcctta tcgcgctccg  3840

gcttcagcaa cagcatccgg atatgcggat cttgcttatt gaggcgggtc ctgaggcggg  3900

agggaaccat acctggtcct ttcacgaaga ggatttaacg ctgaatcagc atcgctggat  3960

agcgccgctt gtggtccatc actggcccga ctaccaggtt cgtttccccc aacgccgtcg  4020

ccatgtgaac agtggctact actgcgtgac ctcccggcat ttcgccggga tactccggca  4080

acagtttgga caacatttat ggctgcatac cgcggtttca gccgttcatg ctgaatcggt  4140

ccagttagcg gatggccgga ttattcatgc cagtacagtg atcgacggac ggggttacac  4200

gcctgattct gcactacgcg taggattcca ggcatttatc ggtcaggagt ggcaactgag  4260

cgcgccgcat ggtttatcgt caccgattat catggatgcg acggtcgatc agcaaaatgg  4320
```

```
ctaccgcttt gtttataccc tgccgctttc cgcaaccgca ctgctgatcg aagacacaca    4380

ctacattgac aaggctaatc ttcaggccga acgggcgcgt cagaacattc gcgattatgc    4440

tgcgcgacag ggttggccgt tacagacgtt gctgcgggaa gaacagggtg cattgcccat    4500

tacgttaacg ggcgataatc gtcagttttg gcaacagcaa ccgcaagcct gtagcggatt    4560

acgcgccggg ctgtttcatc cgacaaccgg ctactcccta ccgctcgcgg tggcgctggc    4620

cgatcgtctc agcgcgctgg atgtgtttac ctcttcctct gttcaccaga cgattgctca    4680

ctttgcccag caacgttggc agcaacaggg gttttttccgc atgctgaatc gcatgttgtt    4740

tttagccgga ccggccgagt cacgctggcg tgtgatgcag cgtttctatg gcttacccga    4800

ggatttgatt gcccgctttt atgcgggaaa actcaccgtg accgatcggc tacgcattct    4860

gagcggcaag ccgcccgttc ccgttttcgc ggcattgcag gcaattatga cgactcatcg    4920

ttgaagagcg actacatgaa accaactacg gtaattggtg cgggctttgg tggcctggca    4980

ctggcaattc gtttacaggc cgcaggtatt cctgtttgc tgcttgagca gcgcgacaag    5040

ccgggtggcc gggcttatgt ttatcaggag cagggcttta cttttgatgc aggccctacc    5100

gttatcaccg atcccagcgc gattgaagaa ctgtttgctc tggccggtaa acagcttaag    5160

gattacgtcg agctgttgcc ggtcacgccg ttttatcgcc tgtgctggga gtccggcaag    5220

gtcttcaatt acgataacga ccaggcccag ttagaagcgc agatacagca gtttaatccg    5280

cgcgatgttg cgggttatcg agcgttcctt gactattcgc gtgccgtatt caatgagggc    5340

tatctgaagc tcggcactgt gcctttttta tcgttcaaag acatgcttcg gccgcgccc    5400

cagttggcaa agctgcaggc atggcgcagc gtttacagta aagttgccgg ctacattgag    5460

gatgagcatc ttcggcaggc gttttctttt cactcgctct tagtggggggg gaatccgttt    5520

gcaacctcgt ccatttatac gctgattcac gcgttagaac gggaatgggg cgtctggttt    5580

ccacgcggtg gaaccggtgc gctggtcaat ggcatgatca agctgtttca ggatctgggc    5640

ggcgaagtcg tgcttaacgc ccgggtcagt catatggaaa ccgttgggga caagattcag    5700

gccgtgcagt tggaagacgg cagacggttt gaaacctgcg cggtggcgtc gaacgctgat    5760

gttgtacata cctatcgcga tctgctgtct cagcatcccg cagccgctaa gcaggcgaaa    5820

aaactgcaat ccaagcgtat gagtaactca ctgtttgtac tctattttgg tctcaaccat    5880

catcacgatc aactcgccca tcataccgtc tgttttgggc cacgctaccg tgaactgatt    5940

cacgaaattt ttaaccatga tggtctggct gaggattttt cgctttattt acacgcacct    6000

tgtgtcacgg atccgtcact ggcaccggaa gggtgcggca gctattatgt gctggcgcct    6060

gttccacact taggcacggc gaacctcgac tgggcggtag aaggaccccg actgcgcgat    6120

cgtatttttg actaccttga gcaacattac atgcctggct tgcgaagcca gttggtgacg    6180

caccgtatgt ttacgccgtt cgatttccgc gacgagctca atgcctggca aggttcggcc    6240

ttctcggttg aacctattct gacccagagc gcctggttcc gaccacataa ccgcgataag    6300

cacattgata atctttatct ggttggcgca ggcacccatc ctggcgcggg cattcccggc    6360
```

```
gtaatcggct cggcgaaggc gacggcaggc ttaatgctgg aggacctgat ttgacgaata    6420

cgtcattact gaatcatgcc gtcgaaacca tggcggttgg ctcgaaaagc tttgcgactg    6480

catcgacgct tttcgacgcc aaaacccgtc gcagcgtgct gatgctttac gcatggtgcc    6540

gccactgcga cgacgtcatt gacgatcaaa cactgggctt tcatgccgac cagccctctt    6600

cgcagatgcc tgagcagcgc ctgcagcagc ttgaaatgaa aacgcgtcag gcctacgccg    6660

gttcgcaaat gcacgagccc gcttttgccg cgtttcagga ggtcgcgatg gcgcatgata    6720

tcgctcccgc ctacgcgttc gaccatctgg aaggttttgc catggatgtg cgcgaaacgc    6780

gctacctgac actggacgat acgctgcgtt attgctatca cgtcgccggt gttgtgggcc    6840

tgatgatggc gcaaattatg ggcgttcgcg ataacgccac gctcgatcgc gcctgcgatc    6900

tcgggctggc tttccagttg accaacattg cgcgtgatat tgtcgacgat gctcaggtgg    6960

gccgctgtta tctgcctgaa agctggctgg aagaggaagg actgacgaaa gcgaattatg    7020

ctgcgccaga aaaccggcag gccttaagcc gtatcgccgg gcgactggta cgggaagcgg    7080

aaccctatta cgtatcatca atggccggtc tggcacaatt acccttacgc tcggcctggg    7140

ccatcgcgac agcgaagcag gtgtaccgta aaattggcgt gaaagttgaa caggccggta    7200

agcaggcctg ggatcatcgc cagtccacgt ccaccgccga aaaattaacg cttttgctga    7260

cggcatccgg tcaggcagtt acttcccgga tgaagacgta tccaccccgt cctgctcatc    7320

tctggcagcg cccgatctag ccgcatgcct ttctctcagc gtcgcctgaa gtttagataa    7380

cggtggcgcg tacagaaaac caaaggacac gcagccctct tttcccctta cagcatgatg    7440

catacggtgg gccatgtata accgtttcag gtagcctttg cgcggtatgt agcggaacgg    7500

ccagcgctgg tgtaccagtc cgtcgtggac cataaaatac agtaaaccat aagcggtcat    7560

gcctgcacca atccactgga gcggccagat tcctgtactg ccgaagtaaa tcagggcaat    7620

cgacacaatg gcgaatacca cggcatagag atcgttaact tcaaatgcgc ctttacgcgg    7680

ttcatgatgt gaaagatgcc agccccaacc ccagccgtgc atgatgtatt tatgtgccag    7740

tgcagcaacc acttccatgc cgaccacggt gacaaacacg atcagggcat tccaaatcca    7800

caacataatt tctcaagggc gaattcgcgg ggatcctcta gagtcgacct gcaggcatgc    7860

aagcttggca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca    7920

acttaatcgc cttgcagcac atcccccttt cgccagctgg cgtaatagcg aagaggcccg    7980

caccgatcgc ccttcccaac agttgcgcag cctgaatggc gaatggcgct gatgtccggc    8040

ggtgcttttg ccgttacgca ccaccccgtc agtagctgaa caggagggac agctgataga    8100

aacagaagcc actggagcac ctcaaaaaca ccatcataca ctaaatcagt aagttggcag    8160

catcacccga cgcactttgc gccgaataaa tacctgtgac ggaagatcac ttcgcagaat    8220

aaataaatcc tggtgtccct gttgataccg ggaagccctg gccaactttt ggcgaaaat    8280

gagacgttga tcggcacgta agaggttcca actttcacca taatgaaata agatcactac    8340

cgggcgtatt ttttgagtta tcgagatttt caggagctaa ggaagctaaa atggagaaaa    8400
```

```
aaatcactgg atataccacc gttgatatat cccaatggca tcgtaaagaa cattttgagg    8460

catttcagtc agttgctcaa tgtacctata accagaccgt tcagctggat attacggcct    8520

ttttaaagac cgtaaagaaa aataagcaca agttttatcc ggcctttatt cacattcttg    8580

cccgcctgat gaatgctcat ccggaattt                                     8609
```